1982

An analysis of price responses to public information: a case study of the USDA corn crop forecasts

Jin Wook Choi
Iowa State University

8221179

Choi, Jin Wook

AN ANALYSIS OF PRICE RESPONSES TO PUBLIC INFORMATION: A CASE STUDY OF THE USDA CORN CROP FORECASTS

Iowa State University PH.D. 1982

University
Microfilms
International 300 N. Zeeb Road, Ann Arbor, MI 48106

An analysis of price responses to public information:

A case study of the USDA corn crop forecasts

by

Jin Wook Choi

A Dissertation Submitted to the

Graduate Faculty in Partial Fulfillment of the

Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Major: Economics

Approved:

Signature was redacted for privacy.

In Charge of Major Work

Signature was redacted for privacy.

For the Major Department

Signature was redacted for privacy.

For the Graduate College

Iowa State University
Ames, Iowa

1982

TABLE OF CONTENTS

LIST OF TABLES

Page

CHAPTER I. INTRODUCTION

General Problem

The concept of perfect information plays a very important role in the theory of a perfectly competitive market. A satisfactory definition of perfect information, however, is never given by economists. Mansfield [32, p. 224], for instance, comes close to defining it to be an accurate knowledge of the future together with the past and present economic and technological data, while Henderson and Quandt [21, p. 105] narrow its meaning to be complete information with respect to the quality and nature of the product and the prevailing price. Cole [10, p. 204], on the other hand, stresses the importance of the time element in defining perfect information by insisting that it should be complete and instant knowledge such that the response to the information on market changes should be virtually instantaneous. Nonetheless, the following remark made by Kamerchen and Valentine [28, p. 149] about perfect information expresses most candidly the current feeling of many economists:

> "perfect" knowledge is, in general, as foolish a goal as "perfect" ignorance . . . no consumer would ever seek out perfect knowledge (even if it would be defined unambiguously, which is also dubious).

And yet, no economist can ignore the importance of information in the decision-making of the market participants. Notwithstanding the problem of defining "perfect information", the economist is still left with the problem of defining "information." Even accepting the

Webster definition of information as

> Something received or obtained through informing as
>
> a. knowledge communicated by others or obtained from investigation, study, or instruction,
>
> b. knowledge of a particular event or situation,
>
> c. facts or figures ready for communication or use as distinguished from those incorporated in a formally organized branch of knowledge,

the puzzle over the distinction among data, information, and knowledge remains. This is because the above definition of information treats data, information, and knowledge as synonymous. However, some economists such as Eisgruber [12, p. 1542], Dunn [11, pp. 19-20], and Bonnen [3, p. 758] claim that "data are not information" because data are symbolic of some phenomena which they are designed to represent while information is a process which imposes form upon and gives meaning to data. Information also differs from knowledge because, in order for information to be such a process, different fields of knowledge should be combined and used. Given this exposition and the business definition of information being "those cues which have the potential to affect managerial decision [66, p. 529]," the meaning of information in economics becomes less obscure. That is, the role of information in economics as in the business disciplines lies in the problem-solving or decision-making purpose. Therefore, information in economics should in a broader sense mean a result of the synthesizing process of data with other knowledge to aid the problem-solving or decision-making purpose.

In summary, without borrowing the words from Hicks [22, pp. 1-11] and Bonnen [3, pp. 753-761], one can readily agree that economics is a science of information-processing because it is especially concerned with the decision-making and with the consequences obtained from the implemented decision, and because the decision somehow made has its basis in some information. Nonetheless, as in the case of describing some abstract concepts in other disciplines (e.g., Hicks [22, p. 106] shows the nebulous circulatory nature in defining "random" in statistics), the often-used term, "information," is not defined to a functional level in economics as it has been argued above. And yet, we economists are to understand and analyze the role and effect of information on the market.

Specific Problem and Objectives

As shown in the previous section, information is an essential element in decision-making. For the grain markets, for instance, the information affecting the market price is so essential that even reports of questionable accuracy were once preferred to none at all [23, pp. 99-102]. However, despite the fact that the complete identification of all possible sets of information and their interrelationships which may affect the prices of grain markets is impossible, let us suppose that any information[1] can be classified

---

[1]No explicit attempt will be made to distinguish among data, information, and knowledge throughout this dissertation. Even though we recognize the inherent differences among them, "information" will be used to represent all three concepts unless the need for specific distinction arises.

into either public or private information, depending upon its source and nature of availability. Public information is then information provided generally by nonprofit organizations at minimum cost or free of charge for its dissemination while private information is information provided generally by profit-motivated organizations at prices over the costs, and thus, its availability is limited. Given this classification, then, the information obtained from the United States Department of Agriculture (USDA)--as long as its distribution is not impaired in any way--are "public." Even though there may be lots of public information other than from the USDA, the traders in the grain markets pay particular attention to the USDA information mainly because it is the major source of relevant information to the grain market traders [23, pp. 99-102].

The public information can affect both the demand for and the supply of grains. For the supply side of the grain market, the following types of public information issued by the USDA may influence the market price: (a) information on estimated yield, planted acreages, projected harvest, export or import conditions, volume of stocks in storage, etc.; (b) the weather-related information such as drought condition, flood, early frost, crop diseases in the grain producing areas, etc.; and (c) the changes in the government policies such as loan rate changes, grain embargo, abandonment of the set-aside or land-retirement programs, etc. [23, pp. 317-334; 4, pp. 22-30]. Thus, it is easy to note that such public information as the USDA corn crop forecasts is only a small subset of an information set which affects the market.

Given this limitation, the specific problem addressed in this dissertation is to analyze the accuracy of the USDA corn crop forecasts in comparison to the actual crop size of a given year and to identify whether these data are processed into information that influences the corn market prices.

Therefore, the objectives of this study are to examine: (a) the accuracy of the USDA corn crop forecasts issued in July through December for the 1930-1977 period; and (b) the effect of these corn crop forecasts upon Iowa corn cash and Chicago corn futures prices. The specific questions embedded in the objective (b) can be more narrowly specified as: (i) how do the size and direction of the USDA corn crop forecast changes from month to month affect the corn market price movements, and (ii) what is the time lag of price adjustments to the USDA corn crop forecast?

Following Chapters

The organization of this dissertation is as follows. Chapter I has introduced the general problem of defining "information" in economics and outlined the specific problem and objectives in analyzing the USDA corn crop forecasts.

Chapter II is devoted to the accuracy analysis of the USDA corn crop forecasts in terms of the differences between these forecasts and their final crop size estimates.

Chapter III develops a theoretical framework to identify those factors, including the USDA corn crop forecasts, which influence the day-to-day price movements and an empirical framework to analyze what impact the USDA corn crop forecasts have on the cash and futures prices of corn. In addition, the adjustment period of price responses to the USDA corn crop forecasts will be studied.

Chapter IV then summarizes the findings and draws conclusions about the USDA corn crop forecasts based on those findings. Also, recommendations will be made for further studies in this area.

CHAPTER II. ACCURACY OF USDA CORN CROP FORECASTS

It is commonly known that many traders in corn markets pay close attention to the USDA corn crop forecast information. Gunnelson, Dobson, and Pamperin [20, p. 639] state:

> A current evaluation of the accuracy of the crop forecasts appears useful since farmers, agribusiness firms, and government agencies make decisions involving billions of dollars annually on the basis of the forecasts, and deficiencies in the forecasts may cause undesired effects on plans and resource allocation.

The impression one receives from this statement is that researchers generally assume that the information obtained from accurate forecast data is less likely to cause undesired effects on plans and resource allocation than the information obtained from inaccurate data. Even though this assumption seems to hold true for many situations, it is believed to be a question under empirical verification which is to be done inter-disciplinarily. Thus, we will not examine here such hypotheses of information processing which seem to be more closely related to psychology than to economics. Rather, this chapter will be devoted to the accuracy analysis of the USDA corn crop forecasts between 1930 and 1977 without intentionally making any inferences on information processing in human decision-making.

Data Set

The Crop Reporting Board (CRB) of the United States Department of Agriculture (USDA) is responsible for collecting and disseminating projected crop estimates in the United States.[1] Even though these estimates are reported in the monthly publication, Crop Production, different branches in the USDA were responsible for their dissemination as the structure of the USDA changed over the years.[2] Nonetheless, the CRB has had the prime responsibility for making the projected crop estimates available.

The CRB is careful to differentiate the terms used in its reports. A "projection," for instance, refers to an expected crop size which may be realized when broad assumptions of crop growing conditions and of the growers' intentions are met. A "forecast" of the crop harvest refers to an expected crop size obtained by examining the maturing crop condition at the time of survey and by assuming that normal growing condition will prevail until the harvest time. An "estimate," on the other hand, refers to the crop size estimated after the crop is fully matured and mostly harvested [58, p. 2]. However, for the purpose of brevity and convenience, the CRB recently uses the terms, "indicated production" or "forecast," to mean all the statistics on the crop size data in general as long as no confusion over the meaning of these data seems to be present.

---

[1] In addition to these production estimates, CRB estimates stocks, inventories, disposition, utilization, and prices of agricultural commodities, and such other items as labor, farm numbers, and fertilizer [58, p. i].

[2] For instance, Bureau of Agricultural Economics was responsible for these crop estimates during 1943-53 period; Agricultural Marketing Service for 1954-60 period; Statistical Reporting Service for 1961-77 period; and Economics and Statistical Crop Service for the 1978-present period.

The corn crop forecasts are the following data: July, August, September, October, and November indicated production estimates. In addition to these, the CRB publishes a final crop size estimate in December of that year or in January of the following year.[1] And yet, this is not the final estimate on the corn crop size of a given crop year.[2] There are two more estimates. The first revised final estimate comes out a year later, after additional revisions on the December final estimate, and the other comes out five years later in a summary bulletin, Field Crops: Estimates By States, after all the final revisions are made.

Therefore, there are eight different estimates of the corn crop size for a given crop year. Taking the revised final estimate which is available five years later to be the true crop size of a given crop year, we evaluate in this chapter the accuracy of the USDA corn crop forecasts published in July, August, September, October, November, and December or January from 1930 till 1977.[3]

---

[1] Up to 1970, the final year-end crop estimates were published in December. Thus, these estimates are often called the December final estimates. However, since 1971, they are being published in January of the following year. These estimates shall also be referred to as the December final estimates for the purpose of the analysis herein.

[2] A crop year for corn begins on October 1, and ends on September 30 of the following year.

[3] The reason why the crop forecasts released after 1978 are not included in this analysis is that the five-year revised final estimates since 1977 are not available as yet.

Before we proceed to analyzing the accuracy of the USDA corn crop data, one final point of our interest is concerned with the missing data on the July corn crop forecasts during the corn blight years from 1971 through 1974. During these years, the USDA refused to present them and delayed its appraisal of the corn crops until August because it did not wish to provide the market participants with uncertain information. That is, the following excerpt from The Des Moines Register on July 12, 1980, shows well the basic objection to the July forecasts in general:

> The July 1 USDA corn forecast usually draws some criticism because the information used to prepare the estimates is compiled before the all-important tasseling-pollination period when the kernels of grain are actually formed on the cob [36, p. 1].

Despite this valid criticism, however, the USDA has continued to provide the July forecasts to the public since 1974.

Previous Analysis

The earliest study on the nature of the USDA crop forecasts was done by Sarle [47] in 1932 when he examined the adequacy and reliability of these data on the basis of their sampling procedures. Thus, he concentrated his effort mainly on the sampling properties of the USDA crop estimates. However, he made an overall appraisal of the historical estimates of the yield-per-acre by comparing them to the census data which was taken several months to a year after the crop was harvested [47, p. 102]. The method used was the simple evaluation of their differences and their correlation coefficients. He found that between 1879 and 1924, the USDA estimates (or official estimates) of corn were within 2 bushels of the census enumeration 50 percent of the time and the correlation coefficient between them was over 0.90. Furthermore, he concluded that

> There is a tendency in the case of corn for both the sample data and the official estimates of yield to be higher than yields derived from census data. This tendency is probably due to the fact that an appreciable proportion of the corn acreage is not harvested for grain, but is used otherwise [47, p. 120].

Even though Sarle's findings may have been accurate for the period of his analysis, the developments in sampling theory and procedures as well as the efficiency gained through technical changes in data management require a new and fresh examination of the USDA crop forecasts. However, unlike Sarle, our concentration of the analysis centers on the accuracy of these data, not on their adequacy, where accuracy here means the magnitude of the difference between the forecast value and the actual value.

Along this line, Clough [9] studied the accuracy of the USDA corn crop forecasts between 1929 and 1950 by comparing the March indicated acreages of corn to the actual acreages harvested, and by comparing the December estimates of production to the estimates made in earlier months. By resorting to a simple regression, he concluded that the March intentions reflected more than 80 percent of the year-to-year variation in corn acreage while the July forecasts accounted for about 60 percent of the variation in corn production. Furthermore, he found that the forecasts in successive months became progressively nearer the December estimates, indicating that the accuracy of the USDA corn crop forecasts increased from month to month.

About the same time, Baker and Paarlberg [1] studied the wheat reports. Based on an accuracy evaluation score and an error reduction score, both of which were inventions of their own, they evaluated the wheat price forecasts, wheat production forecasts, and wheat carry-over forecasts of the USDA. The basic accuracy evaluation score is the average of the arbitrary scores given to various forecasts according to the actual changes that occurred. That is, a score of 100 is given when the forecast direction of change is the same as the actual change, 50 if the direction of the forecast is either increasing or decreasing while the actual outcome has no change, 25 if the forecast is for no change when the actual outcome changes, and 0 if the forecast and actual change are in opposite directions to each other. Therefore, they assert that an accuracy evaluation score ranges from 0 to 100 "with 50 representing the score that would theoretically be obtained if random forecasts were made over a long period of time" [1, p. 105-6]. Based on this scoring system, the winter wheat production forecasts between 1938 and 1951 received an overall accuracy score of 78 and appeared commendable to them. As for the error reduction score, they computed it by taking account of the variation in the forecast series. That is, the average annual variation in wheat production minus the average percentage error in monthly forecasts, divided by the former average annual variation, yields the error reduction score in a percentage when the resulting number is multiplied by 100. The winter wheat production forecasts received a score of 51, indicating that a reduction in error of 51 percent was provided by the winter wheat production forecasts.

Gunnelson, Dobson, and Pamperin [20] analyzed the accuracy of the USDA crop production forecasts for barley, corn, oats, potatoes, soybeans, spring and winter wheat for the 1929-1970 period, based on an established statistical method. By employing Theil's R statistic (or revision ratio), the accuracy of the USDA corn crop forecasts was noted to be consistently improved over the years. And in checking the systematic error (or bias) in the USDA forecasts, it was found that about 60 percent of the first corn crop forecasts during the 1959-1968 period tended to underestimate the actual crop size, which was taken to be the December estimate of that year.

Smith [49], on the other hand, concentrated on the accuracy of the USDA soybean crop forecasts from 1950 to 1971 in an effort to evaluate the improved soybean crop information acquired by remote-sensing technology. By analyzing the mean absolute deviations of monthly soybean crop forecasts from the one-year revised final estimates, he found that the USDA soybean crop forecasts were relatively accurate and its accuracy had steadily been improved.

When Pearson and Houck [40] studied the accuracy of the USDA corn and soybean forecasts between 1963 and 1975, they relied on the graphical trace of the monthly forecasts within each crop year and computed the maximum and average differences between the monthly production forecasts and the December estimates. As for the corn production forecasts, they found congruent results with other studies, in that (a) no systematic biases seem to occur, and (b) a definite trend toward more accurate forecasts exists as the season progresses toward the harvest time.

There is one analysis on the accuracy of the USDA hog farrowing intentions statistics between September of 1959 and March of 1973, done by Thompson [54]. He uses the Theil's R statistic to compare the accuracy improvement of one forecast over the preceding ones and the mean-square-error statistic to detect a systematic bias in the series of forecasts. He finds the same conclusion as found in the analyses of the grains production forecasts. That is, there is no systematic bias in the hog farrowing statistics and their accuracy improves over the years.

Present Analysis

It can thus be concluded from the previous section that the previous analyses of the USDA corn crop forecasts were concerned with the question of accuracy, accuracy improvement, and systematic bias. Interestingly enough, however, the techniques used in these analyses except the regression analysis are of nonparametric nature because there is no explicit assumption on the distribution of the forecast errors, however they may have been defined. Even though the use of nonparametric tests for accuracy is not wrong, it is believed that the proper knowledge on the distribution of forecast errors can provide us with a better insight in selecting appropriate statistical tools for testing the accuracy, accuracy improvement, and systematic biases.

Therefore, we shall first study the nature of the USDA corn crop forecast data for the 1930-1977 period by testing the normality of the monthly forecast errors, which are defined in this analysis as the difference between the corn crop forecasts of each month and the five-year revised final estimate for a given crop year. Then, based on the results found from the normality test, we shall choose appropriate statistical tools for the tests of accuracy, accuracy improvement, and systematical bias. That is, if the monthly forecast errors are found to be normally independently distributed,[1] then we can employ Hotelling's T^2 statistic and regression analysis to test the accuracy and accuracy improvement in addition to some nonparametric tools. The test of hypothesized ordering can, for instance, be used in testing the specific order of accuracy which is measured by the absolute magnitude of the mean of the forecast errors in each month.

<u>Normality of forecast errors</u>

First, the forecast errors are defined in this analysis as the

[1]The test of normality used in this study does not test the independence of samples from year to year. The independence of samples are assumed herein.

differences (or deviations) between the monthly corn crop forecasts from July to December (or January) and the five-year revised final estimates:

$$FE_{my} \equiv MF_{my} - FRFE_{y} \tag{1}$$

where FE_{my} is the forecast error of the m-th month's forecast in the y-th year, MF_{my} is the USDA corn crop forecast issued in the m-th month of the y-th year, and $FRFE_y$ is the five-year revised final estimate of corn production in the y-th year.[1] Given these FE_{my}'s, then, we have many possible statistical tools to test their normality. Among them, we have a goodness-of-fit test through a chi-square statistic, the Kolmogorov-Smirnov D-statistic, and the Shapiro-Wilk W-statistic.

The most often used test of normality is the chi-square goodness-of-fit criterion, which is based on the relative frequency of particular observations. This chi-square test is believed to be a nonspecific test, in that the test criterion does not distinguish any particular type of departure, such as a noticeable skewness or flatness, from normality. Therefore, even though this chi-square test renders insight to the normality of given observations,

---

[1] The year y should refer to the crop year rather than a calendar year. However, for simplicity and convenience, it is stated here that the corn crop forecasts (or estimates) of the calendar year y means the crop size of the crop year which begins October 1 of the y-th year and ends September 30 of the (y+1)-th year.

supplementary tests of skewness and flatness are often needed for a proper identification of normality. Further, the conclusion obtained from it often depends upon the choice of a proper class limit into which the observations are arbitrarily classified and one must take note of the conflict between the continuity of a normal distribution curve and its approximation for discrete observations in computing the observed frequency in each class.[1] Regardless of these shortcomings, this chi-square test is often used in testing the normality, and proper procedures for its application are well-described in many elementary statistics textbooks [e.g., 50, pp. 84-90].

While the chi-square test uses the squared values of the differences between the observed and expected relative frequency density functions, the Kolmogorov-Smirnov test uses the absolute values of the differences between the observed and expected cumulative probability distribution functions. Since its concise theoretical development and practical examples of numerous applications are available in many statistics textbooks such as Gibbons [18, pp. 56-77], the assumptions and characteristics of this Kolmogorov-Smirnov D statistic will be briefly examined here. First, this test assumes that the observations are obtained from a random process and that the needed parameters in describing

---

[1]This is an inevitable problem inherent in any other statistical tests of normality because observed data can not accurately be measured or recorded on a continuous scale.

the hypothesized distribution--e.g., the mean and variance of a normal distribution--are known. Further, the D statistic is defined to be the largest absolute difference between the observed and hypothesized cumulative probability distribution functions. Therefore, if the largest value is small, it can be logically inferred that all deviations must be small. Thus, the test criterion is to reject the null hypothesis if D is too large.

The main characteristics of this D test are as follows. First, D statistic is especially appropriate when the sample distribution to be tested is continuous, thus implying an infinite number of classes (or groups). Therefore, this D test is exact for a hypothesis test concerning a continuous population with all needed parameters specified while it is conservative for discrete cases because no general adjustment tool such as a correction factor in the chi-square test is available. Secondly, the use of two-sided D statistic can determine a minimum sample size because the D statistic indicates the maximum absolute deviation between the observed and expected cumulative density functions. Third, the D statistic can be extended to a general distribution test for two or more independent samples, which is sometimes called the Birnbaum-Hall test [50, p. 259].

Even though the use of the chi-square goodness-of-fit or the Kolmogorov-Smirnov D test can be of service to our purpose, the test of normality of forecast errors employed here is, however, the Shapiro-Wilk W statistic. As initially explored by Shapiro and Wilk [48], the W statistic is obtained by dividing the square of an appropriate linear combination of the sample order statistics by the estimate of variance. The underlying concept of the W statistic is based on the fact that a linear transformation of a normal distribution is also a normal distribution. That is, if the y_i's are a sample from a normal distribution, then each y_i can be expressed as

$$y_i = \mu + \sigma x_i \qquad \text{for } i = 1, 2, \ldots, n \tag{2}$$

where $x_1 \leq x_2 \leq \ldots \leq x_n$ are an ordered random sample of size n from a standard normal distribution, and μ and σ are respectively the unknown mean and standard deviation of y_i's.

Let $M = (M_1, M_2, \ldots, M_n)$ denote the vector of expected values of standard normal order statistics (i.e., x_i's), and let $V = (V_{ij})$ be the corresponding n x n covariance matrix. Then, applying the generalized least-squares theorem, the best linear unbiased estimates of μ and σ are respectively found to be $\bar{y}$ and a linear combination of y_i's. That is,

$$\hat{\mu} = \bar{y} \quad \text{and} \quad \hat{\sigma} = \frac{M'V^{-1}y}{M'V^{-1}M} \tag{3}$$

where $\hat{\mu}$ and $\hat{\sigma}$ are respectively the estimates of μ and σ.

Denoting

$$S^2 = \sum_{i=1}^{n} (y_i - \bar{y})^2 \tag{4}$$

which is the usual symmetric unbiased estimate of $(n-1)\ \sigma^2$, the W test statistic for normality is defined by

$$W = \frac{R^4\hat{\sigma}^2}{C^2S^2} = \frac{b^2}{S^2} = \frac{(\sum_{i=1}^{n} a_i y_i)^2}{\sum_{i=1}^{n} (y_i-\bar{y})^2} \tag{5}$$

where
$$R^2 = M'V^{-1}M \tag{6a}$$

$$C^2 = M'V^{-1}V^{-1}M \tag{6b}$$

$$a' = (a_1, a_2, \ldots, a_n) = \frac{M'V^{-1}}{(M'V^{-1}V^{-1}M)^{1/2}} \tag{6c}$$

and $\quad b = R^2\hat{\sigma}/C$

Shapiro and Wilk [48, p. 593] state that "b is, up to the normalizing constant C, the best linear unbiased estimate of the slope of a linear regression of the ordered observations, y_i, on the expected values, M_i, of the standard normal order statistics. The constant C is so defined that the linear coefficients are normalized." Noting that the coefficients, a_i's, are the normalized best linear unbiased coefficients from the table in [48, p. 603-605], Shapiro and Wilk [48, p. 593] show that the maximum value of W is 1 because the numerator and denominator of W are both estimating the same quantity, namely σ^2, if a sample is indeed from a normal population.

To compute the W statistic, we first order the observations to obtain an ordered sample of size n such as $y_1 \leq y_2 \leq \cdots \leq y_n$, and compute its sum of squares of deviations as S^2 in the equation (4). Then, we compute the sum of the linear combination of the differences between appropriately paired sample order statistics.[1] That is

$$b = \sum_{i=1}^{k} a_{n-i+1}\,(y_{n-i+1} - y_i) \tag{7}$$

If n is even (i.e., n=2k), then the equation (7) implies

$$b = a_n(y_n - y_1) + a_{n-1}\,(y_{n-1} - y_2) + \cdots$$

$$+\, a_{k+2}\,(y_{k+2} - y_{k-1}) + a_{k+1}\,(y_{k+1} - y_k) \tag{7a}$$

If n is odd (i.e., n=2k+1), the elimination of the sample median of y_{k+1} means that the equation (7) will become

$$b = a_n(y_n - y_1) + a_{n-1}\,(y_{n-1} - y_2) + \cdots$$

$$+\, a_{k+3}\,(y_{k+3} - y_{k-1}) + a_{k+2}\,(y_{k+2} - y_k) \tag{7b}$$

Thus, given the values of S^2 and b, we can obtain the W statistic as the ratio of b^2 over S^2.

The properties of this W statistic are: (a) it is scale and origin invariant, (b) the maximum value is 1 and the minimum value is $n\,a_1^2/(n-1)$,[2] (c) it offers good power against a large class of

---

[1]The reason for paired sample order statistics is that the normal distribution is symmetric; thus, the weights associated with each statistic are the same.

[2]When n=44 as in the case of July forecast errors, the minimum value is 0.1534 given $a_1 = 0.3872$. When n=48 as in all other cases, the minimum value is 0.1466 given $a_1 = 0.3789$.

alternative hypotheses even for small samples (i.e., $n < 20$), and (d) it is sensitive to outliers, in that the nature of overall configuration of the sample is taken account of. The main drawback of this W statistic is, however, that for large sample sizes (i.e., $n > 50$) the tabulation of the coefficients, a_i's, in the numerator of the statistic may be difficult, cumbersome, and inexact. When the W statistic is computed as above for each set of monthly forecast errors from equation (1), we find the results in Table 1. In addition to this W statistic, Table 1 contains three estimates of location parameters (i.e., the mode, the median, and the mean), three values of dispersion measure (i.e., the range, the mean absolute deviation, and the standard deviation), and the measures of skewness and kurtosis in terms of the third and fourth moments about the mean.[1]

Before interpreting the results in Table 1, we present a brief description of each measure in it as follows. Even though the mode is not a good measure of central tendency because it often depends on the arbitrary grouping of the data, it approximately identifies the value or interval that has the highest frequency in a distribution. The median, which is sometimes called the middle value or the 50-th percentile, is defined to be the value below which half the values in the sample fall. Thus, the median splits the

---

[1]Many elementary statistics books [e.g., 18, 50, and 63] have a good description on these measures of statistical importance.

Table 1. W and other relevant statistics in testing the normality of FE_{my}'s in each month, 1930-1977

Statistics	July	Aug.	Sept.	Oct.	Nov.	Dec.
	Monthly forecast errors					
n	44	48	48	48	48	48
			(unit: 1000 bushels)			
MODE[a]	0	0	0	0	0	0
MEDIAN[b]	-11,875.5	-65,434.5	-77,963	-36,283.5	-9,723.5	-2,505
MEAN	24,415.9	-46,551.7	-47,185	-28,244.1	-5,345.4	-1,254.7
RANGE	1,228,759	1,172,467	748,735	663,163	448,029	311,389
MAD	251,563.7	183,515.7	135,733.9	105,737.9	79,483.1	61,078.4
S	318,039	235,221	175,890	140,441	103,924	76,605.4
			(no unit of measurement)			
SKEW	0.5857*	0.2040	-0.0901	-0.4082	0.0135	0.2403
KURT	-0.1175	0.1037	-0.0406	0.6934	-0.0128	-0.4227
W	0.9479	0.9861	0.9561	0.9616	0.9752	0.9722
PROB<W	0.074	0.917	0.142	0.252	0.541	0.464

[a]The value of MODE used here is the mid-point of an interval which covers -100 million to +100 million bushels of forecast errors, because this interval has the highest frequency in all months.

[b]The value of MEDIAN is obtained by dividing the two mid-values in the observations by 2.

*Means "significant at a 5 percent significance level." Otherwise, no significance is detected at 5 percent level.

observations into two halves. The mean, on the other hand, is the average value of the observations and is equivalent to the center of gravity or the balancing point of the observed data.

As for the measures of dispersion, the range simply measures the difference between the largest and the smallest observations. The mean absolute deviation (MAD), on the other hand, is the average of the absolute values of the deviations between the observed values and their mean. Even though the mathematical problem of differentiating the absolute value function limits the wide use of this measure, it provides a simple and easy assessment of the dispersion. The most common measure of dispersion, however, is the standard deviation (S) which is defined as the positive value of the square root of the variance.

The measures of skewness (SKEW) and kurtosis (KURT) render us more insights into the distribution. A measure of the skewness in a distribution is the third moment about the mean. To make this measure independent of the scale on which the data are recorded, it is divided by the cube of the standard deviation. Snedecor and Cochran [50, p. 86] state that if the sample comes from a normal population, this measure of skewness is approximately distributed with the mean zero and the standard deviation of $\sqrt{6/n}$.[1] If this measure is significantly larger than its standard deviation,

---

[1]When $n=44$, $\sqrt{6/n} = 0.3693$; when $n=48$, $\sqrt{6/n} = 0.3536$. However, these values of the standard deviation(s) are found in the table by Snedecor and Cochran [50, p. 552]. There, we find $S = 0.3418$ when $n=45$; $S = 0.3264$ when $n=50$.

positive skewness (or right-skewedness) is present. If the measure is significantly smaller than its standard deviation, negative skewness (or left-skewedness) is present. That is, if the low values of the data are bunched close to the mean but high values extend far above the mean, this measure will be positive and will show a right-skewedness because the large positive contributions to the third moment about the mean when the observed value is larger than the mean will predominate over the smaller negative contribution to the third moment when the observed value is smaller than the mean. Therefore, we will note in general that if the mode is less than the median which is in turn less than the mean, the distribution will be skewed to the right and the measure of skewness will be positive.

The measure of kurtosis is the fourth moment of the sample about its mean divided by the squared value of its variance. According to Snedecor and Cochran [50, pp. 86-87], this ratio has the value of 3 for a normal distribution. It is thus concluded that if the measure of kurtosis exceeds 3, then the sample distribution shows more peakedness than a normal distribution. Values less than 3, on the other hand, result from distributions that have a flatter top than the normal. Thus, the usual measure of kurtosis is obtained by subtracting 3 from the fourth moment about the mean divided by the squared value of the variance.[1]

---

[1]The distribution of the measure of kurtosis does not approach the normal distribution until the sample size is larger than 1000 [50, p. 88]. Geary [17, p. 295] developed an alternative method for a smaller sample size but it was not used here.

Equipped with this background information on the statistical meanings of various statistics presented in Table 1, we can now readily draw the following conclusions about the distribution of the monthly forecast errors of the USDA corn crop forecasts. First, the measures of central tendency, being negative, show that the monthly USDA corn crop forecasts generally underestimate the five-year revised final estimates. However, the decreasing trends observed in the absolute magnitudes of MEDIAN and MEAN suggest that the accuracy of the forecasts improves over the months. Even though this hypothesis of accuracy improvement will be statistically tested in the next section, we have a similar evidence of accuracy improvement from the measures of dispersion in Table 1. That is, the largest values of the range, MAD, and the standard deviation are found in the month of July and the smallest values of these statistics are found in December. Further, these values are decreasing during the intervening months, which suggests that the density of the USDA monthly forecasts around the five-year revised final estimate increases as the reporting month approaches harvest.

As for the skewness and flatness of the monthly forecast errors, we detect no significant departures from normality except that the July forecast errors display some skewness to the right at the 5 percent significance level. This finding is understandable in light of the positive mean value and a relatively closer-to-zero median. However, the negative values of kurtosis for the July, September, November, and December forecast errors suggest some flatness in their distributions but are not significant at the 5 percent significance level. Further, the negative skewness observed in September and October forecast errors shows that these distributions may be skewed to the left but the skewness is negligible at the 5 percent significance level. Thus, this examination of the skewness and flatness of the monthly forecast errors suggest that they have a normal distribution, which is in fact the conclusion drawn from the W statistics.

From the bottom line of Table 1, the prob-values[1] associated with the W statistics are all larger than 0.05, suggesting that the null hypothesis of a normal distribution of the monthly forecast errors should not be rejected at the 5 percent significance level. The case which comes closest to the rejection of the null hypothesis is the distribution of the July forecast errors, where the prob-value is 0.074. This observation of a low prob-value for July is not so surprising because we have observed that July forecast errors had a significant skewness to the right.[2] This detection of a

---

[1]The prob-value is defined as the probability of the sample value being as extreme as the value we actually observed, assuming the null hypothesis to be true [63, p. 190]. Thus, a judgment criterion is: "reject H_0 if and only if the prob-value is less than or equal to a chosen significance level [63, p. 195]."

[2]This significant skewness was detected at the 5 percent but not at the 1 percent significance level. If independent tests of 12 null hypotheses were made at the 1 percent significance level, the probability of no type I error made in these 12 tests would be $(0.99)^{12}$=0.886 and that of some error would be $(1-0.99^{12})$=0.114. For a detailed explanation, see [63, p. 301].

significant skewness at the 5 percent significance level, however, may be attributed to the compounded type I error (i.e., the error of rejecting H_o when H_o is true) which is associated with the repetitive tests of independent null hypotheses. That is, since we have tested the significance of 6 skewness measures and 6 kurtosis measures at the 5 percent significance level (or at the 95 percent confidence level), we note that the probability of no type I error at all for these 12 tests is $(0.95)^{12}=0.54$, which means that the probability of some type I error is as high as $(1-0.95^{12})=0.46$. Therefore, in consideration of this probability for falsely rejecting the null hypothesis when it is true and because of the nonsignificance of the W statistic, the significant skewness in July forecast is taken to be a type I error and does not lead to rejection of the normality of its distribution. As for the other monthly forecast errors, the W statistics detect no significant departure from a normal distribution. Thus, we can conclude that all monthly forecast errors have a normal distribution at the 5 percent significance level.

Hotelling's T^2

The purpose of this section is to analyze the accuracy of the monthly forecasts using the findings of the previous section. The major conclusions were: (a) the forecast errors seem to be smaller in their magnitudes as the reporting months progress, and (b) the monthly forecast errors are distributed normally. However, we have not yet drawn any statistical conclusion about the accuracy of each monthly forecast.

Since we define accuracy in terms of the magnitude of the difference between the forecast and actual value (i.e., the forecast error), a monthly forecast series is said to be accurate if the mean of the forecast errors is statistically not different from zero. However, even if we reject the null hypothesis of a zero mean of particular monthly forecast errors at a specified level of significance, it does not imply that that monthly crop forecast is useless. In other words, if we say, for instance, that the mean of the August forecast errors is statistically different from zero (i.e., inaccurate), it does not mean that the August crop forecasts present an inaccurate picture of the crop condition prevailing as of August 1 (and thus are useless). The term, accuracy, used in this analysis only means that degree of exactness (or closeness) of the August crop forecasts in comparison to the five-year revised final estimate which is taken to be the true crop size of that year.

The nature of the crop forecasts is such that the estimates are taken from a same object over different periods. Therefore, if we assume that the forecasts of different years are independent and that the forecasts can be decomposed additively, we can write the forecast of the m-th month in year y (i.e., MF_{my}) as

$$MF_{my} = \mu_y + \mu_m + e_{my} \tag{8}$$

where μ_y is a general crop level common to all monthly crop forecasts in year y, μ_m is the effect unique to the month m, and e_{my} is the random disturbance. Further, if we assume that μ_y is the true crop size of year y (i.e., μ_y = $FRFE_y$), then based on the equation (1), we can express the forecast error of month m in year y as

$$FE_{my} = \mu_m + e_{my} \qquad \text{for } \begin{matrix} m = 1, 2, \ldots, 6 \\ y = 1, 2, \ldots, n \end{matrix} \tag{9}$$

If the monthly effect, μ_m, is fixed and the vector of variates $[e_{1y}, e_{2y}, \ldots, e_{6y}]$ has the multivariate normal distribution[1] with mean vector

$$E[e_{1y}, e_{2y}, \ldots, e_{6y}] = [0, 0, \ldots, 0] \tag{10}$$

and covariance matrix

$$\Sigma = E\left(\begin{bmatrix} e_{1y} \\ e_{2y} \\ \cdot \\ \cdot \\ \cdot \\ e_{6y} \end{bmatrix} [e_{1y}, e_{2y}, \ldots, e_{6y}]\right) \tag{11}$$

then, we can employ the Hotelling's T^2 statistic to test the null hypothesis of all monthly effects being zero versus the alternative hypothesis of not all means being zero.[2]

---

[1]Since we have found that the forecast errors, FE_{my}'s, are normally distributed from the previous section, e_{my}'s are also normally distributed given fixed μ_m's.

[2]The underlying assumptions for the test of this null hypothesis are the same as the fixed-effects model of the Analysis of Variance. However, the interest of the fixed-effects model is to test the null hypothesis of equal effects rather than all effects being zero.

The necessary statistics in computing the Hotelling's T^2 are as follows.[1] Since we have assumed that

$$e_{my} \sim \text{NID}\ (\bar{0}, \Sigma) \tag{12}$$

where $\bar{0}$ is a 6 x 1 zero matrix and Σ is a 6 x 6 variance-covariance matrix, we observe from the equation (9) that

$$FE_{my} \sim \text{NID}\ (\mu, \Sigma) \tag{13}$$

where μ is a 6 x 1 matrix of monthly effects; i.e., $\mu' = [\mu_1\ \mu_2\ \ldots\ \mu_6]$. The unbiased estimates of μ and Σ from the sample of size n are then obtained by

$$\hat{\mu}_m = \frac{1}{n} \sum_{y=1}^{n} FE_{my} \tag{14}$$

$$S_{mk} = \frac{1}{n-1} \sum_{y=1}^{n} (FE_{my} - \hat{\mu}_m)(FE_{ky} - \hat{\mu}_k) \tag{15}$$

where $\hat{\mu}_m$ and S_{mk} are respectively the sample mean of the m-th month's forecast errors and the sample covariance between the forecast errors of months m and k. Denoting the estimates μ and Σ to be $\hat{\mu}$ and S respectively, we can express $\hat{\mu}$ and S in a matrix notation as follows:

$$\hat{\mu} = \begin{bmatrix} \hat{\mu}_1 \\ \hat{\mu}_2 \\ \cdot \\ \cdot \\ \cdot \\ \hat{\mu}_6 \end{bmatrix} \quad \text{and } S = \begin{bmatrix} S_{11} & S_{12} & \cdots & S_{16} \\ S_{21} & S_{22} & \cdots & S_{26} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ S_{61} & S_{62} & \cdots & S_{66} \end{bmatrix} \tag{16}$$

---

[1] For a detailed discussion, see Morrison [35].

Given these statistics, we are now to test the null hypothesis of all monthly effects being zero against the alternative hypothesis that not all monthly effects are zero. That is,

$$\begin{aligned} H_o&: \ \mu = \bar{0} \\ H_a&: \ \mu \neq \bar{0} \end{aligned} \tag{17}$$

When the null hypothesis is true, the quantity T^2 (which is called Hotelling's T^2) has the F-distribution with P=6 and (n-P)=42 degrees of freedom:

$$T^2 = n(\hat{\mu}-\bar{0}) \quad S^{-1}(\hat{\mu}-\bar{0}) \tag{18}$$

$$F = \frac{n-P}{P(n-1)} T^2 \tag{19}$$

Since T^2 is a scalar, it can be easily noted from the equation (18) that departures of $\hat{\mu}$ from zero can only increase the value of T^2. Therefore, the larger is T^2, the larger is the chance of rejecting the null hypothesis. However, a more specific decision rule for the test of a null hypothesis at a α percent significance level is:

$$\text{Accept } H_o: \ \mu=\bar{0}$$

$$\text{if } \frac{n-P}{P(n-1)} T^2 \leq F_{\alpha:\ P,n-P} \tag{20}$$

and reject H_o otherwise.

When T^2 is computed by the equation (18), based on the sample statistics summarized in Table 2, we obtain T^2 to be 8.6418 and the

Table 2. The sample mean and coded sample variance-covariance matrix of the monthly forecast errors, 1930-1977 (unit: 1,000 bushels)

		Coded variance-covariance[a]					
Month	Mean	July	Aug.	Sept.	Oct.	Nov.	Dec.
July	24,415.9	101.15	49.66	20.40	9.63	3.33	-1.67
Aug.	-46,551.7	49.66	55.33	32.14	19.76	10.03	4.58
Sept.	-47,185.0	20.40	32.14	30.94	22.45	12.44	7.12
Oct.	-28,244.1	9.63	19.76	22.45	19.72	12.91	7.69
Nov.	-5,345.4	3.33	10.03	12.44	12.91	10.80	6.89
Dec.	-1,254.7	-1.67	4.58	7.12	7.69	6.89	5.87

[a]The coded variance-covariance matrix should be multiplied by 10^9 to obtain the actual variance-covariance matrix.

computed F statistic, F_c, to be 1.2871, based on the equation (19). We find that the computed F statistic of 1.2871 is much smaller than the tabulated F statistic of 2.32 with (6,42) degrees of freedom at the 5 percent significance level. Consequently, the null hypothesis that all monthly means are not different from zero is not rejected. That is, the monthly forecasts are said to be accurate, in that the differences between their means and the five-year revised final estimates are statistically nonsignificant.

Nonparametric L

Even though the monthly forecasts are found to be statistically accurate in the previous section from the joint test of the null hypothesis of zero means for all forecast errors, we have observed in Table 1 that the medians and means of the forecast errors tend to approach zero and that the measures of dispersion tend to be smaller as the reporting month approaches December. This qualitative observation is congruent to common intuition which suggests that the crop forecasts in the earlier months (e.g., July) are less accurate than forecasts in later months (e.g., December)[1] due to a larger uncertainty on crop-growing conditions in, say, July than in December. The accuracy analysis in the previous section has not directly dealt with this hypothesis of accuracy

[1]It would be interesting to study the hypothesis that the forecasts of the earlier years, when compared to the recent years, would be less accurate due to lack of data-gathering and computing efficiencies, and the changes in the farming structure in general. However, this hypothesis will not be examined in this dissertation.

improvement over the reporting months; rather, it simply concluded that each of the monthly forecast errors had a mean zero.

In this section, we shall therefore consider the question of whether the accuracy of the USDA corn crop forecasts improves over the reporting months by the following two approaches. First, we qualitatively interpret the frequency distribution of forecast errors in Table 3 in addition to the statistics in Table 1 to draw conclusions about accuracy improvement. The second approach is based on the nonparametric L test for a hypothesized ordering.

The conclusion drawn from the qualitative observation on the monthly forecast errors in the top part of Table 1 and in Table 3 is that the accuracy does improve over the reporting months because there is a general tendency for the magnitude of the forecast errors in later months to be closer to zero. Table 3 shows that the forecast errors have greater dispersion in the earlier months and concentrate more closely around zero in the later months. This fact is numerically expressed and verified in the measures of dispersion in Table 1. It is, however, interesting to note that the July, November, and December forecast errors are quite evenly distributed around the mean zero (i.e., from the fact that the numbers of under- and over-estimations are relatively equal) while there is a definite under-estimation tendency in August, September, and October. However, there is no reason to expect such a pattern. That is, since the corn crop forecasts made by the USDA only

Table 3. Frequency of monthly forecast errors falling in each interval for the 1930-1977 period

Interval (mill. bu.)	Frequency of forecast errors						
	July	Aug.	Sept.	Oct.	Nov.	Dec.	
-600 and below		1					1
-500 to -599							0
-400 to -499	5	1	1	1			8
-300 to -399	1	6	3	2			12
-200 to -299	4	5	3	2	2		16
-100 to -199	7	8	11	8	5	5	44
0 to -99	6	10	12	17	18	20	83
0 to 99	5	4	8	9	15	17	58
100 to 199	3	5	5	6	6	6	31
200 to 299	6	4	5	3	2		20
300 to 399	2	2					4
400 to 499	1	1					2
500 to 599		1					1
600 and above	4						4
Underestimates	23	31	30	30	25	25	164
Overestimates	21	17	18	18	23	23	120
Total	44	48	48	48	48	48	284

represent the expected crop size as of the first day of the announcement month, only random influence from the changes in the crop growing condition are expected. Thus, we can reasonably expect that the under- and over-estimations of the five-year revised final estimates should be equally likely. This hypothesis can be stated as in the following set of hypotheses (21) and tested by the normal approximation method of a binomial distribution.[1] That is,

$$\begin{aligned} H_o &: \ P = 1/2 \\ H_a &: \ P \neq 1/2 \end{aligned} \qquad (21)$$

where P is the probability of under-estimation. Then, we can compute the Z statistic as

$$Z = (|r-np|-1/2)/\sqrt{np(1-p)} \qquad (22)$$

where r is the observed frequency of under-estimation, n is the total number of observations, and 1/2 is used as the correction for continuity. Thus, we obtain the computed Z statistic, Z_c, as

---

[1]According to Snedecor and Cochran [50, p. 211-213], the two-tailed Z-test using the Z statistic (obtained by the normal approximation method of a binomial distribution as computed in the equation (22)) is equivalent to the chi-square test using chi-square with 1 degree of freedom, except that the formula for chi-square has no correction for continuity. The exact relationship is that chi-square is the squared value of Z in the equation (22) without the correction for continuity. However, for one-sided alternatives, the Z-test is preferred to the chi-square test which is basically two-sided.

$$Z_c = (|166-284(0.5)|-0.5)/\sqrt{284(0.5)(0.5)}$$

$$= 2.5516 \tag{23}$$

Z_c is much larger than the tabulated Z of 1.96 at the 5 percent significance level. We, therefore, reject the null hypothesis in favor of the alternative hypothesis at the 5 percent significance level.[1] That is, the under- and over-estimations of the five-year revised final estimates are not equally likely in the series of the USDA corn crop forecasts. However, this unequal tendency of the under- and over-estimations observed in the data in most attributable to the August, September, and October corn crop forecasts, during which the under-estimations occurred approximately by a two-to-one ratio. In spite of this under-estimation tendency, we note that approximately one-half of the forecast errors fall in the intervals between -99 million to 99 million bushels and the frequencies observed in these intervals steadily increase from July to December as it is shown in Table 3. Thus, we qualitatively conclude for now that the accuracy improves over the reporting months and next examine this hypothesis through a nonparametric test.

---

[1]When the alternative hypothesis in (21) is stated as $P > 1/2$, we should compare Z_c to the tabulated Z value at the 5 percent significance level which is about 1.65 for this one-sided hypothesis test. In this case, we also reject the null hypothesis in favor of the alternative.

The nonparametric test for the hypothesis of accuracy improvement in the USDA corn crop forecasts is the L test proposed by Page [39]. Since the L test examines the null hypothesis of equal means against an alternative hypothesis of ordered means, the predetermination of the expected ordering of the means to be tested is an essential element of this method. We choose the order of the means as in the set of hypotheses in (24), based on the belief that the mean of July forecast errors will be the largest, and that of December forecast errors will be the smallest, and the in-between months will have the means of an intermediate magnitude. That is, we are to test

$$H_o: \ \mu_1 = \mu_2 = \cdots = \mu_6$$

$$H_a: \ \mu_1 > \mu_2 > \mu_3 > \mu_4 > \mu_5 > \mu_6 \tag{24}$$

where the subscripts 1 through 6 denote the months from July to December respectively, and μ denotes the mean of monthly forecast errors.

Page [39] suggests the following computational steps in using the L test for the above types of a monotonic hypothesis. First, we set up the monthly forecast errors into a two-way table of P (= 6) columns (= months) and n (= 48) rows (= years). Second, since there are 6 groups of monthly forecast errors, we rank the monthly forecast errors for each year from 1 to 6, giving the largest value 6 to the smallest monthly forecast error in absolute magnitude. Third, we sum the ranks in each column, so finding $\sum_{y=1}^{n} X_{my}$ where X_{my} is the value of rank for the m-th monthly forecast errors of the y-th year. Fourth, we multiply each such sum of ranks by the expected rank for that same column. That is, we compute the values of $\sum_{y=1}^{n} X_{my}Y_m$ where Y_m is the value of the expected rank for the m-th monthly forecast error as specified in the alternative hypothesis. Therefore, Y_m will be 1 for the month of July and 6 for December. Fifth, we sum all such products to find the computed L statistic, L_c. That is,

$$L_c = \sum_{m=1}^{P} \sum_{y=1}^{n} X_{my}Y_m \tag{25}$$

The necessary information to compute this L_c is presented in Table 4. By adding the elements in the bottom row of Table 4, we obtain L_c = 4,013 which is much larger than 3,697, the tabulated value of L at the 0.001 significance level. Therefore, we reject the null hypothesis at this 0.001 level and conclude that the predicted and the observed rankings in the monthly mean-errors are in agreement. That is, the accuracy of the USDA corn crop forecasts does improve over the reporting months. This statistically verifies the qualitative observation made earlier on month-to-month accuracy improvement.

However, this finding of accuracy improvement through the L test seems to contradict the earlier finding of equal means through the Hotelling's T^2 statistic. That is, the latter showed that no means of monthly forecast errors differ from zero at the 5 percent

Table 4. Frequency of observed rank in each month and the needed statistics for computing L, 1930-1977

Rank	Monthly frequency July	Aug.	Sept.	Oct.	Nov.	Dec.
1	23	13	6	3	1	0
2	5	20	12	5	4	0
3	5	4	21	8	5	5
4	4	3	2	19	9	11
5	5	3	3	6	18	13
6	2	5	4	7	11	19
Sum of ranks ($\sum_{y=1}^{n} X_{my}$)	101	122	140	185	216	238
Expected rank (Y_m)	1	2	3	4	5	6
$\sum_{y=1}^{n} X_{my} Y_m$	101	244	420	740	1080	1428

significance level; thus, no improvement in accuracy over the reporting months was suspected. This apparent conflict can be explained by a close scrutiny into the nature of the F test based on the Hotelling's T^2 and that of the L test. Their main difference lies on the fact that the F test used in Hotelling's T^2 examines the cardinal strength of the data while the L test examines the ordinal strength. The cardinal strength as the term is used here refers to the magnitude of the actual values as they are observed in their original units of measurement while the cardinal strength refers to a nominal (or ordinal) scale given to the observed values.[1] Therefore, such sample parameters as the mean and variance represent the cardinal strength of the observations while the rank or sign assigned to the observations represent the original strength of the data. T^2 uses the mean and variance-covariance matrices and tests the null hypothesis of equal cardinal strength while L uses the value of the ranks and tests the null hypothesis of equal ordinal strength. For the specific conflict observed in this analysis, we note that T^2 depends upon the magnitudes of the sample means and the sample variance-covariance as noted in the equation (18). That is, either an increase of

[1]It can be roughly stated here that the statistical inference techniques of parametric nature test the cardinal strength while those of nonparametric nature test the ordinal strength of the data. For a better discussion of this topic, see [18, pp. 22-30].

numerical magnitudes in the mean vector or a decrease of values in the variance-covariance matrix will result in a larger value of computed T^2 which implies that the probability of rejecting the null hypothesis of all means being zero increases. However, these changes in the mean and/or variance-covariance do not affect the significance level of the L test as long as the rank of the observations remains unchanged. This point can be better explained by the following hypothetical example. If the July forecast errors are observed to be -100 million and 100 million bushels, and the December ones -10 million and 10 million bushels. Then, their respective means are both zeroes. Consequently, there is a high probability of accepting the null hypothesis of equal means. However, we observe that the forecast errors of July are larger than those of December; thus, there exists a definite order in their magnitudes. Of course, we expect that not all forecast errors of July would be greater than those of December in our data set. However, the L statistic basically tests this relationship of ordered observations while T^2 statistic provides a test for equal magnitudes of the means of monthly forecast errors. Therefore, the tests of no order and of the equal magnitudes of means are both compatible, but not identical. Thus, Snedecor and Cochran [50, p. 132] notes that the rank tests are about 95 percent efficient in large normal samples but slightly more efficient in small normal samples relative to the t-test in comparing equal means. In light of this consideration, therefore, we can conclude that the conflict of conclusions drawn from the Hotelling's T^2 and the L test does not nullify one conclusion or the other. That is, the null hypothesis of all means being zero, tested and accepted within the Hotelling's T^2 criterion, shows that the magnitudes of the means of forecast errors are statistically equivalent to zero while the conclusion obtained by the L test states that forecast errors decrease from July to December.

Regression

This section presents findings on the accuracy and accuracy improvement of the USDA corn crop forecasts through the regression analysis. We first estimate simple regressions of the following form:

$$FRFE_y = \alpha_m + \beta_m \cdot MF_{my} + e_{my} \tag{26}$$

where $FRFE_y$ is the five-year revised final estimate of the crop size in year y; MF_{my} is the USDA corn crop forecast in the month m of year y; α_m and β_m are, respectively, the intercept and the slope coefficient associated with month m; and e_{my} is the random disturbance term with the property of NID $(0, \sigma_e^2)$. Thus, if we are to say that the forecast of month m is an accurate estimate of the final crop size, then α_m should be zero and β_m should be one. That is, we are to test the following set of hypotheses:

$$\begin{aligned} &H_o: \ \alpha_m = 0 \text{ and } \beta_m = 1 \qquad \text{for } m = 1, 2, \ldots, 6 \\ &H_a: \ \alpha_m \neq 0 \text{ or } \beta_m \neq 1 \end{aligned} \tag{27}$$

The appropriate test criterion for this hypothesis is obtained by the F statistic which is given by Ostle and Mensing [38, p. 174]. We first define that A_m and b_m are respectively the estimated regression coefficients of α_m and β_m and that S_e^2 is the residual mean squares of the regression (or the estimate of σ_e^2). Then, we note that

$$Q = \sum_{y=1}^{n} [(A_m + b_m\, MF_{my}) - (\alpha_m + \beta_m\, MF_{my})]^2/\sigma_e^2 \tag{28}$$

is distributed as a chi-square with 2 degrees of freedom and that $(n-2)\ S_e^2/\sigma_e^2$ is distributed as a chi-square with (n-2) degrees of freedom. Therefore, it can be seen that the computed F statistic,

$$F_c = \frac{Q/2}{[(n-2)S_e^2/\sigma_e^2]/(n-2)} = \frac{Q\sigma_e^2}{2S_e^2} \tag{29}$$

is distributed as F with 2 and (n-2) degrees of freedom. That is, if F_c is less than the tabulated F value with 2 and (n-2) degrees of freedom at a 5 percent significance level,[1] we fail to reject the null hypothesis; otherwise, we reject the null hypothesis in favor of the alternative hypothesis.

After the coefficients were estimated by an ordinary least squares method for the regression equation (26) which is named as

---

[1]The tabulated F values with (2,46) degrees of freedom at the 5 percent and the 1 percent significance levels are respectively 3.20 and 5.10, and F with (2,42) degrees of freedom at the 5 percent significance level is 3.22 [50, p. 626].

Model I, F values were computed based on the information shown in Table 1 of Appendix. These computed F values, presented in Table 5, lead us to conclude that for all months except September, the intercept term is not different from zero and the regression coefficient is equal to one at the 5 percent significance level. The rejection of the null hypothesis (27) for September was somewhat surprising but it was not confirmed when the null hypothesis was tested at the 1 percent significance level. Therefore, the regression without the intercept term which is named as Model II is generally concluded to be a better choice for describing the relationship between the final crop size and the monthly forecasts. That is, given the relationship

$$FRFE_y = \gamma_m \, MF_{my} + e_{my} \tag{30}$$

whose estimation results are also presented in Table 5, we tested the null hypothesis that

$$H_o: \ \gamma_m = 1 \qquad \text{for all m's}$$

$$H_a: \ \gamma_m \neq 1 \tag{31}$$

The appropriate test criterion is based on the t statistic which is computed as

$$t_c = \frac{C_m - 1}{S_{Cm}} \tag{32}$$

where t_c is the computed t value, C_m is the estimate of γ_m, and S_{Cm}

Table 5. Estimated regression coefficients of the monthly USDA corn crop forecasts with $FRFE_y$ as the dependent variable, 1930-1977: Models I and II

Monthly forecast	Model I				Model II		
	A_m	b_m	F	R^2	C_m	S_e^2 [b]	R^2
July	-23,845.2 (169,721.9)[a]	0.9998 (0.0473)	0.1279	0.9141	0.9935 (0.0134)	10.1196	0.9923
Aug.	-15,014.7 (114,522.7)	1.0149 (0.0324)	0.8217	0.9590	1.0108 (0.0094)	4.8366	0.9963
Sept.	-58,657.8 (79,412.5)	1.0296 (0.0225)	3.2527*	0.9803	1.0137* (0.0065)	2.3479	0.9982
Oct.	-99,564.1 (68,064.4)	1.0354 (0.0192)	3.1616	0.9858	1.0085 (0.0056)	1.7604	0.9987
Nov.	-92,710.0 (52,212.1)	1.0265 (0.0146)	2.1310	0.9916	1.0017 (0.0044)	1.0691	0.9992
Dec.	-64,055.2 (38,307.7)	1.0172 (0.0107)	0.6872	0.9954	1.0001 (0.0032)	0.5817	0.9996

[a]The values in parentheses represent the standard errors of each estimate.

[b]S_e^2 is the residual mean square of the regression, which should be decoded by multiplying by 10^{10}.

*Denotes the rejection of the null hypothesis at the 5 percent significance level but not at the 1 percent significance level.

is the standard error of C_m. Therefore, a statistical conclusion at the δ percent significance level can be drawn as

If $|t_c| \leq t_{(\delta/2,n-1)}$, fail to reject H_o.

Otherwise, reject H_o. (33)

We note from the results under Model II in Table 5 that all C_m's except that of September are equal to one at the 5 percent significance level.[1] As for September, the rejection of the null hypothesis (27) was attributable to the regression coefficient, b_m, not being equal to one, which was confirmed by the rejection of the null hypothesis (31) at the 5 percent significance level. However, both null hypotheses could not be rejected at the 1 percent significance level. In light of these findings, we can thus conclude and confirm the following facts which have already been discovered in the previous sections.

First, the monthly forecasts (except that of September) are in general accurate estimates of the final crop size because the relationship (30) with γ_m being equal to one seems to be true.

---

[1]The tabulated t values with 45 degrees of freedom at the 5 percent and the 1 percent significance levels are respectively 2.014 and 2.690, and the t value with 40 degrees of freedom at the 5 percent significance level is 2.021 for a two-tailed test [50, p. 549].

Second, all C_m's (as well as b_m's) except for July are greater than one, indicating that monthly forecasts in August through December have the tendency to under-estimate. This under-estimation tendency seems to be strongest in the September forecasts because the regression coefficient, C_m, is statistically not equal to one and is largest in September. Third, the values of R^2 progressively increase in both models, indicating that the accuracy of the monthly forecasts do improve over the reporting months. We note that R^2 is the ratio of the explained variation over the total variation. For the cast of a simple regression as used in this analysis

$$R^2 = (\sum_{i=1}^{n} X_i Y_i)^2 / (\sum_{i=1}^{n} X_i^2 \sum_{i=1}^{n} Y_i^2) \tag{33a}$$

where Y_i is the independent variable and X_i is the regressor. When a regression was run with an intercept, X_i and Y_i used in the formula for R^2 represent deviations from their respective means. In a regression without an intercept, X_i and Y_i are actual values.

Even though the previous regression analysis provided us with a good description of the relationships between the final crop size and the monthly forecasts, the following extension of the regression analysis for the purpose of predicting the USDA corn crop forecasts can be of value. That is, how can we best use the known USDA forecasts to predict what the USDA forecast of the next month of the final crop size will be? This question of practical importance will be examined next.

First, we ask how the final crop size can best be predicted from the known USDA monthly corn crop forecasts. That is, how can the July, August, and September forecasts best be used to predict the final crop size? It seems at a first glance that the following form of a regression may suggest a possible answer:

$$FRFE_y = \alpha_k + \sum_{m=1}^{k} \beta_m MF_{my} + e_{ky} \qquad \text{for } k = 1, 2, \ldots, 6 \qquad (34)$$

where α_k is the intercept of the k-th regression equation, β_m is the regression coefficient associated with the forecast of month m, and e_{ky} is the random disturbance term with the property of NID $(0, \sigma_e^2)$. When this regression equation (34) was estimated, we found that the intercept terms were not significantly different from zero and that the USDA forecasts of those months prior to the most recent month seemed to be of no value in predicting the final crop size. Therefore, the hypothesis that the USDA corn crop forecast of the most recent month alone can provide sufficient information in predicting the final crop size was tested on the basis of the following regression equation:

$$FRFE_y = \sum_{m=1}^{k} \gamma_m MF_{my} + e_{ky} \qquad \text{for } k = 1, 2, \ldots, 6 \qquad (35)$$

The hypothesis to be tested can be formally stated as

$$H_o: \ \gamma_1 = \gamma_2 = \cdots = \gamma_{k-1} = 0 \qquad (36)$$

$$H_a: \ \text{not } H_o$$

i.e., in terms of the model in equation (35), the hypothesis is a comparison of the two models in (37)

$$H_o: \ FRFE_y = \gamma_k \ MF_{ky} \tag{37}$$

$$H_a: \ FRFE_y = \gamma_1 \ MF_{1y} + \gamma_2 \ MF_{2y} + \dots + \gamma_{k-1} \ MF_{k-1,y} + \gamma_k \ MF_{ky}$$

where in the model in H_a at least one γ_m ($m = 1, 2, \dots, k-1$) is not zero.

The appropriate test procedure is based on partitioning correctly the total sum of squares (TSS) into the regression sum of squares (RSS) and the error sum of squares (ESS).[1] Denoting the regression sum of squares due to including the independent variables MF_{1y}, MF_{2y}, ..., MF_{ky} as RSS ($\gamma_1, \gamma_2, \dots, \gamma_k$), we can express the total sum of squares in the following two ways. When the model in H_o of (37) is run, we get

$$TSS = RSS\ (\gamma_k) + ESS_o \tag{38}$$

When the model in H_a of (37) is run, we get

$$TSS = RSS\ (\gamma_1, \gamma_2, \dots, \gamma_k) + ESS_a \tag{39}$$

where TSS's in (38) and (39) are equal. The test statistic for the hypothesis (36) is computed as

[1]For the usual case where the intercept term is included, we would note that TSS = RSS + ESS can be specifically stated as $\Sigma(Y_i-\bar{Y})^2 = \Sigma(\hat{Y}_i-\bar{Y})^2 + \Sigma(Y_i-\hat{Y}_i)^2$ where Y_i is the observed dependent variable, $\bar{Y}$ is its mean, and $\hat{Y}_i$ is the estimated value of Y_i [63, p. 338].

$$F_c = \frac{[RSS(\gamma_1, \gamma_2, \ldots, \gamma_k) - RSS(\gamma_k)]/(k-1)}{ESS_a/(n-k)} \qquad (40)$$

which follows an F distribution with (k-1,n-k) degrees of freedom. The decision criterion is as follows:

If $F_c < F_{\alpha(k-1,n-k)}$, fail to reject H_o.

Otherwise, reject H_o. (41)

The sums of squares needed to compute F statistics are presented in Table 2 of the Appendix, and the resulting F statistics are shown in Table 6 along with the estimated regression coefficients. We observe that none of the F statistics in Table 6 are greater than the tabulated F value at the 5 percent significance level, indicating that the null hypothesis stated in (36) or (37) should not be rejected. This result is also shown by the simple t tests of the coefficients as noted in Table 6. Furthermore, the values of R^2 in Table 6, computed from the multiple regression equation (35), are almost identical to those R^2's in Table 5 obtained from the simple regression equation (30), indicating that the past monthly forecasts do not effectively reduce the unexplained variation of the model. Furthermore, the residual mean squares of the two models in (37) are nearly equal as shown in Tables 5 and 6. Therefore, we can conclude that the simple regression equation (30) and its corresponding values of coefficients are sufficient in predicting the final corn crop

Table 6. Regression coefficients of the monthly corn crop forecasts in predicting the final crop size and the F statistics for selecting the appropriate prediction equations, 1930-1977

Current month	Regressors, coded S_e^2, and R^2								
	July	Aug.	Sept.	Oct.	Nov.	Dec.	S_e^{2a}	R^2	F^b
July	0.9935* (0.0134)[c]						10.1196	0.9923	0
Aug.	-0.0149 (0.1491)	1.0259* (0.1514)					4.9506	0.9963	0.0099
Sept.	0.0986 (0.1054)	-0.1557 (0.2048)	1.0696* (0.1589				2.4085	0.9982	0.4591
Oct.	0.1074 (0.0877)	0.1265 (0.1821)	-0.8349 (0.4534)	1.6047* (0.3655)			1.6658	0.9988	1.8138
Nov.	0.0575 (0.0674)	-0.0449 (0.1423)	0.5681 (0.4305)	-1.2769* (0.5963)	1.6954* (0.3102)		0.9675	0.9993	2.1301
Dec.	0.0718 (0.0503)	0.0111 (0.1065)	0.0993 (0.3313)	-0.2512 (0.4798)	-0.2232 (0.4097)	1.2913* (0.2276)	0.5376	0.9996	1.7052

[a]S_e^2 is the residual mean squares of the regression which should be decoded by multiplying them by 10^{10}.

[b]The tabulated F value with (5,38) degrees of freedom at the 5 percent significance level is 2.46.

[c]The value in parentheses represent the standard errors of each estimated regression coefficient.

*Denotes that the coefficient is significantly different from zero at the 5 percent significance level when t test is employed.

size. That is, the USDA corn crop forecast of the most recent month is most relevant information for the purpose of predicting the final crop size.

Conclusions

In this study, the five-year revised final estimate was taken to be the true measure of crop size. Therefore, the forecast errors were defined as the differences between the monthly USDA corn crop forecasts and their five-year revised final estimates. The accuracy was then measured by the absolute magnitude of these differences. That is, the smaller the absolute value of a forecast error was, the more accurate the monthly corn crop forecast was.

Since the use of the Shapiro-Wilk W statistics enabled us to conclude that the forecast errors for each month from July to December were distributed normally, an additional assumption of the independent distribution of forecast errors in different years allowed us to use an F-test based on Hotelling's T^2 statistic to test the null hypothesis that the means of the monthly forecast errors are all zero. This null hypothesis was not rejected at the 5 percent significance level and it was concluded that monthly USDA corn crop forecasts were accurate and their means were not different from the five-year revised final estimates.

However, common intuition and observation of the forecast data suggested that the series of monthly forecasts within each reporting year improve over the reporting months. This hypothesis of accuracy improvement was then tested by a nonparametric L test and accepted at the 0.1 percent significance level. That is, the null hypothesis of equal means was rejected in favor of an alternative hypothesis of ordered means which stated that the magnitudes of forecast errors become progressively smaller from July to December. Thus, the increasing accuracy of the monthly series of forecasts was verified.

The regression of the five-year revised final estimates on two or more monthly crop forecasts showed that the only significant monthly forecasts were the most recent forecast. That is, if the forecasts of July, August, and September were known, for instance, the coefficient associated with the September forecast was the only meaningful value in explaining the size of the five-year revised final estimate. Thus, this finding was taken to mean that the accuracy improves over the reporting months. Furthermore, this fact implied that in forecasting the true crop size of a given crop year, the most recent month's forecast was its best estimate. However, as it was verified by a chi-square test, the August, September, and October forecasts under-estimated the five-year revised final crop estimates twice as frequently as they over-estimated. This fact was shown in the regression coefficients associated with these months' forecasts by being slightly larger than one, which must be taken into account in forecasting the final crop size.

In summary, we noted that: (a) the overall forecast errors were not different from zero implying that the USDA corn crop forecasts were unbiased; (b) the accuracy improves over the reporting months; (c) the under-estimation of the final crop size occurs twice as frequently as the over-estimation during August, September, and October; and (d) for a forecasting purpose, the most recent crop forecast is the best estimate of the final crop size.

CHAPTER III. PRICE RESPONSES TO USDA CORN CROP FORECASTS

This chapter will examine whether or not the USDA corn crop forecasts have any impact on the corn cash and futures markets.

Review of previous studies done on information analysis will be presented first. Then, by concentrating our attention to the corn market, we shall develop theoretical and empirical models on the basis of reservation demand theory and supply-of-storage theory. Finally, the impact of the USDA corn crop forecasts on the corn prices will be estimated and interpreted.

We must realize, however, that an overall assessment of the role played by the USDA corn crop forecast information in the corn markets can never be exact. This is because how one transforms and interprets the announced crop data into useable information and how one integrates or incorporates this information into his decision-making is not yet fully known. Even if one fully knows this decision-making process, still the problem of correctly separating an effect of one piece of information from all other information on the variable of our interest may remain insurmountable.

We assume that (a) the figure on expected crop size becomes information to corn traders when the Crop Reporting Board ends its secrecy by announcing its estimate, and (b) the announcement of the corn crop forecast is the only source of systematic disturbance in the market during the period of our analysis; namely, five trading days before and after the announcement. Any other disturbances that occur during this period are assumed to have randomly distributed effects on corn prices. Based on these assumptions, the analysis presented in this chapter explores a theoretical and empirical estimation procedure to tackle the problem of identifying and measuring the impact of the USDA corn crop information on the corn prices.

Introduction: Economics on Information

Before we construct theoretical and empirical models to analyze the impact of USDA corn crop information on corn markets, we shall briefly review in this section what the economists have accomplished in analyzing the role of information in the market.

Even though such fundamental questions as what information is, how it is to be measured, and what constitutes an improvement or an increase in information still remain unresolved [7, p. 347], the economists, since the early 1960s, have become particularly aware of the need to investigate the value and role of information in the market. Their efforts can largely be divided into three areas of information analysis.

The first area of work deals with the search for information, which is exclusively studied under the topic of the information search theory. This theory treats information concerning the market prices as an economic good [e.g., 16, 45, and 52].

The second area of information analysis deals with the use of information. The rational expectations hypothesis and the efficient market hypothesis are two prominent treatments of the way people, i.e., the decision-makers, use information in making their trading decisions. The rational expectations hypothesis [e.g., 29, 37, and 46] states that people do not waste information in forming their expectations about the future event, and thus the expected result is statistically congruent to the actual outcome.[1] On the other hand, the efficient market hypothesis as summarized by Fama [14] states that the market is efficient if all available information is fully reflected in the prices. Thus, the implication of this hypothesis is that no one can consistently make a positive profit in an efficient market because the traders already have taken into account the influence of the factors, i.e., the information, that may affect the market price; that is, the information is fully reflected in the price. Thus, we can readily conclude that these hypotheses are different from the information search theory, in that they are concerned mainly with the rational or efficient use of information.

The third area of information analysis does not yet have any theoretical foundation, for its main interest lies in examining the question of whether or not a certain piece of information or event has altered the pre-existing conditions of, say, a corn market. That is, it analyzes the impact of information on the

---

[1]Muth [37, p. 333] claims, for example, that "the rational expectations hypothesis states that, in aggregate, the expected price is an unbiased predictor of the actual price."

market by identifying the direction and magnitude of a change in a variable or in the structure of the economy which has presumably occurred after the receipt of a given piece of information (or the occurrence of a given event). Therefore, this third type of information analysis differs significantly from those which deal with the search for and the use of information.

Given these three types of information analyses, we can easily conclude that the search for and the use of information do not concern us much in analyzing the importance of the USDA corn crop forecast announcement in the market. Rather, we must concentrate our attention on the problem of identifying and measuring the impact of information and the rapidity of price responses to the information. We shall explore the relevant conceptual approaches in analyzing the impact of information as follows.

Generally, there are two different approaches. The first approach is somewhat complex, in that an economic model is first constructed to predict the values of the variable of our interest (e.g., the corn price) which would have occurred in the absence of new information (e.g., the USDA corn crop forecast in some month). Then, the predicted corn price is compared to the observed corn price to assess the impact of the announcement of the USDA corn crop forecasts. The underlying reason for this type of complex approach is that other factors which affect the determination of a corn price may have changed over time, independently from the receipt of information of our interest; and these changes should have been correctly taken into account in evaluating the impact of given information. Even though this approach has the power to potentially delineate the impact of a specific piece of information from others, as was done by Reid [44] for the wage control program in the United States during the 1960-1978 period, the success of this type of analysis depends crucially on the nature and power of the economic model constructed for the purpose of predicting the variable of our interest.

The other kind of conceptual approach to the information impact analysis assumes that the underlying economic conditions do not change over time. This approach is therefore particularly appealing if one is to investigate the impact of a given piece of information during a very short period such as a day or a week when the underlying economic conditions are believed to be reasonably stable. This simpler approach directly compares the magnitudes of the variable (e.g., the corn price) before and after the occurrence of an event such as the announcement of the USDA corn crop forecast.

Of these two approaches to analyzing the impact of the USDA corn crop information on the corn market, we prefer the simpler approach of comparing the magnitudes of a variable observed before and after a given event (or information) over the complex approach of constructing an economic model to predict the values of the variable in the absence of an event (or information). We shall now review the literature, which used the simpler approach, to analyze the impact of information.

The work done by Pearson and Houck [40] aroused economists' interest in analyzing the impact of the USDA reports on grains and livestock production. Pearson and Houck used a nonparametric chi-square test to examine the hypothesis of an inverse relationship between the changes in the USDA corn crop forecasts and the corresponding changes in the daily cash prices. Their conclusion was that the USDA corn crop forecasts had an impact on the market because forecast changes and prices had an inverse relationship for corn, soybean, and spring wheat, but not for winter wheat, during the 1963-1975 period.

Since then, Gorham [19] ran a regression of a percentage change in prices on a percentage change in forecasts for soybeans, wheat, and corn, and found that only corn demonstrated a statistically significant relationship between the price and the forecast. He concluded that the private market must have anticipated the changes in soybean and wheat forecasts well and the change in corn forecasts poorly. Hoffman [24], on the other hand, analyzed the impact of the quarterly livestock reports on cattle and hog prices by regressing the price differences between the periods before and after the release of USDA livestock reports on the percentage change in an appropriate quantity variable such as cattle on feed and sows farrowing. He found that, on the average, the prices before and after livestock reports are not significantly different. However, for specific reports, such as the percentage change in placements of cattle on feed, sows farrowing, or marketing intentions, the cash market seemed to respond while the futures market did not. This suggested to him that the futures markets for cattle and hogs were more efficient than the cash markets.[1]

Similar regression analysis was used to study live hog futures prices' response to the sows farrowing information [33], the cucumber price response to the marketing order [61], and beef demand in response to promotion programs [43].

Even though these studies have tackled the problem of empirically analyzing the impact of information mainly on the market-clearing prices, they have not given much attention to the adjustment response of the price to the given information. Thus, in the remaining sections of this chapter, we will first present a theoretical analysis to examine the impact of information on the corn market, and then an empirical analysis.

Theoretical Analysis

This section is devoted to examining the nature of price determination in the grain markets in relation to newly available crop

---

[1]This suggested conclusion is objectionable on the ground that the efficient market hypothesis deals with the use of information, not with the impact of information. That is, the significant regression coefficient associated with the reported quantity variables when the cash prices were used as a dependent variable means, in fact, that the traders in the cash market were more responsive to the new information contained in the report, not less responsive. Thus, the cash market may be more efficient in the sense that it takes full account of new information by being more responsive to it.

information. First, we shall use the concept of reservation demand to identify the factors that influence the equilibrium price, and then the supply of storage theory will be used to study intertemporal price differences.

<u>Reservation demand</u>

The price is determined by the demand and supply conditions at any given point in time. That is, if we are to know how a market-clearing price of corn is determined, for instance, it is necessary to know what factors affect the underlying conditions for demand and supply of corn. This has been the subject of much research in agricultural economics [e.g., 55, pp. 32-43 and pp. 80-91; 15; and 60].

Those factors that are generally believed to affect the quantity demanded of corn are its own price, the prices of substitutes and complements, the number of livestock, the prices of livestock and related products, the income of the demanders, their tastes and preferences, and expectations about the future market conditions. However, the identification of those factors that affect the supply of corn is not so straightforward and deserves close scrutiny based on two main characteristics of corn.

The first characteristic is that corn is produced once a year in a large lump sum, it is highly storable over a relatively long period, and its consumption is continuously made from the stored stocks throughout the nonproduction period. Thus, we note that the quantity of corn supplied in any period is only a fraction of the quantity produced in a crop year, and the rest is presumably in storage after the harvest. Therefore, the quantity of corn supplied is largely a function of a storage decision based on the size of crop harvested and stored, and the cost of storage which can be measured in terms of the current and expected prices of corn. The quantity of corn produced, on the other hand, depends largely upon such factors as the expected price at harvest time, the input prices, the planted acreages, the crop growing conditions in relation to weather, crop disease, and insect infestation, etc. Therefore, we note that the decisions to produce and to supply corn in the market are closely related in the long-run, but not so in the short-run.

Another reason that the total quantity of corn produced is not the same as the total quantity supplied (or sold) in the corn market for a given year is attributable to the fact that a large portion of the quantity produced is consumed by its producers.[1] That is, the producer of corn, mainly the farmer, is not only a supplier but also a consumer of corn produced.[2] This dual role of

---

[1]During the 1976-80 period, about 38 percent of total production of corn for grain was consumed on farms where produced (59).

[2]We distinguish herein between a producer and a supplier, and between a demander and a consumer because a producer can be a consumer but not a demander of the commodity in a market. That is, a producer need not enter into the market as a demander to consume the needed amount. He simply consumes what he produces.

a producer, especially in a storable grains market, has long been recognized as reservation demand by economists, e.g., Ezekiel [13], Breimyer [5], and Pesek [41]. The reservation demand is described as follows:

> Reservation demand relates to the supply function of the very short run. It is not output response but a market-release function, concerning the rate at which a stock of a good is released into market channels. Reservation demand is both a supply and demand function, for it is a supplier's own withholding demand determining the quantities supplied on a daily or weekly market [5, p. 685].

What this suggests, then, is that a producer has a market within his firm such that he is both a demander and a supplier to himself.[1] Thus, we can describe a market mechanism for a producer by accounting identities as follows:

$$SS_t = SST_{t-1} + SH_t \tag{42}$$

$$SD_t = SST_t + SC_t + SQ_t \tag{43}$$

$$SS_t = SD_t \tag{44}$$

[1] In economics textbooks, a producer is traditionally assumed to be a supplier and a consumer, a demander. However, for a storable commodity where the concept of reservation demand plays an important role, we must not treat the terms equally. That is, a producer and a consumer are different from a supplier and a demander, respectively, in that the latter terms within the context of this analysis refer to those who exchange the ownership of the commodity in the established market at a price agreed while the former terms refer to those who transact in the within-firm market where no specific transfer of the ownership takes place and no price is actually quoted for the transaction. Further, a consumer refers to a nonproducer of the commodity.

where (the first initial S denotes the variables related to a producer). SS_t, SD_t, and SC_t, respectively, denote the quantities supplied to himself, demanded from himself, and consumed by himself during period t; SH_t and SQ_t are the quantities harvested and sold in the market during period t; and SST_t is the stock level at the end of period t. Thus, the equation (42) states that the quantity supplied to the producer by himself is the sum of the stock held at the end of the previous period and the quantity harvested during the current period t. The equation (43) which represents the reservation demand states that the quantity demanded by the producer is the sum of his stock level at the end of the current period and the quantities consumed and sold during period t. The equation (44) is an accounting identity which shows that the quantities demanded and supplied by the producer should be equal.

By the same token, we can construct definitional equations for corn movements within a consumer unit. That is, a consumer has a role to supply corn to himself out of his current stock and from the purchased quantity, and a role to demand corn from himself for his consumption need and for his new stock holdings. We describe a market within a consumer as follows:

$$DS_t = DST_{t-1} + DQ_t \tag{45a}$$

$$DD_t = DST_t + DC_t + DX_t \tag{45b}$$

$$DS_t = DD_t \tag{46}$$

where (the first initial D denotes the variables related to a consumer) DS_t, DD_t, and DC_t, respectively, denote the quantities supplied to himself, demanded, and consumed by himself during period t; DQ_t and DX_t are, respectively, the quantities bought in the market and exported to foreign countries;[1] and DST_t is the stock level at the end of period t. Therefore, the equation (45a) states that the total quantity he can supply to himself during period t is the stock carried over from the previous period, t-1, plus what he purchases from the market during period t. The equation (45b) states that the total quantity he demands from himself during period t is the sum of his stock level at the end of period t, the quantity he consumes, and the quantity he exports during period t. The equation (46) then shows that the total quantity supplied to himself must be identical to the total quantity demanded by himself during period t.

Given these sets of equations (42) through (46), we can identify what quantities are supplied and demanded during period t in terms of the marketing conditions within each of the consumer and the producer markets. That is, solving the equations in (42) through (44) for SQ_t and the equations in (45a) through (46) for DQ_t, we find

$$SQ_t = SST_{t-1} + SH_t - (SST_t + SC_t) \tag{47}$$

and

---

[1]Even though a producer can be engaged in an exporting activity, we assume here that a consumer alone exports.

$$DQ_t = DST_t + DC_t + DX_t - DST_{t-1} \tag{48}$$

From these equations (47) and (48),[1] we can better identify the factors which affect the quantities supplied (or sold) and demanded (or bought) in the market during period t by examining the determinants of each component in SQ_t and DQ_t. Even though we may treat SST_{t-1}, DST_{t-1}, and SH_t as fixed because they are assumed to be largely a function of past market conditions, we find one significant departure from the traditional specification of the demand and supply equations, in that the quantity supplied (SQ_t) depends on how much the producer is to consume (i.e., SC_t) and to hold as inventory (i.e., SST_t) and that the quantity demanded (DQ_t) includes the demander's decision on how much stock to hold (i.e., DST_t). Therefore, a market equilibrium price for a highly storable commodity such as corn should be a function of the consumption needs of the supplier and demander and of their storage decisions. However, the consumption decisions are not independent from the storage decisions, in that large consumption means a smaller stock to be carried into the next period. The supply of storage theory, which will be discussed in the next section, explains the inter-temporal price difference as a function of the expected inventory behavior.

---

[1] For a market to exist, SQ_t and DQ_t must be positive. In the absence of corn production in period t, this implies that $SST_{t-1} > SST_t$ but there need not be such a clear relationship between DST_{t-1} and DST_t. This partly explains the importance of DST_{t-1} in affecting the market demand during period t, and suggests that the consumer also has a form of a reservation demand as long as $DST_{t-1} \neq 0$. In other words, DQ_t can be negative for an individual consumer, but in aggregate DQ_t must be positive.

<u>Supply of storage</u>

The supply-of-storage theory, as it was initially conceived by Working [64, 65] and later developed by Brennan [6], Telser [53], Weymar [62], and others,[1] renders an explanation of inter-temporal price differences[2] in storable commodity markets where the quantity supplied in a given period is dependent upon the level of stocks in storage, rather than upon the level of production. That is, the commodity under consideration is occasionally produced in a lump sum and is storable over a long period to meet continuous demand over the nonproduction periods. Thus, the supply decision on a storable commodity, say corn, is synonymous with its storage decision. Essentially, the supply-of-storage theory views the inter-temporal price difference to be the price of storage, which is a function of the inventory level.

We shall explore this theory in depth within the framework of the basic model presented in the previous section. Note that the equation (47) represents the quantity supplied and the equation (48) the quantity demanded. Thus, in equilibrium, the quantities demanded and supplied must be equal. By equating the equations (47) and (48), we can better differentiate those components which

---

[1]Stein [51], Beckmann [2], and Pliska [42], for example, present modified versions of this theory.

[2]An inter-temporal price difference refers to the price difference between any two periods. Thus, roughly speaking, it is synonymous with a price spread.

affect the supply from those affecting the consumption as follows:

$$SST_{t-1} - SST_t + SH_t + DST_{t-1} - DST_t = DC_t + DX_t + SC_t \qquad (44)$$

Denoting $ST_t = SST_t + DST_t$ and $C_t = DC_t + SC_t$, we can rewrite the equation (44) as

$$ST_{t-1} - ST_t + SH_t = C_t + DX_t \qquad (50)$$

We note from this equation (50) that the total quantity disappearing through consumption and export is equal to the change in stock levels between periods t-1 and t plus the quantity harvested. Thus, the equation (50), being identical to the underlying assumption of Brennan's model on the supply of storage, can directly be used to examine the inter-temporal price differences. However, we will interpret the equation (50) in a different manner from Brennan.

Brennan [6, pp. 51-52] assumes that consumption depends only upon the price of the same period and writes the demand function in period t as

$$P_t = f_t\ (C_t + DX_t),\ \frac{\partial f_t}{\partial (C_t + DX_t)} < 0 \qquad (51a)$$

or

$$P_t = f_t\ (ST_{t-1} - ST_t + SH_t),$$

$$\frac{\partial f_t}{\partial ST_{t-1}} < 0,\ \frac{\partial f_t}{\partial ST_t} > 0,\ \frac{\partial f_t}{\partial SH_t} < 0 \qquad (51b)$$

where P_t is the price in period t and f_t denotes that the functional relationship of demand may change over time. He then defines the demand for storage from periods t to t+1 to be

$$P_{t+1} - P_t = f_{t+1} (C_{t+1} + DX_{t+1}) - f_t (C_t + DX_t) \qquad (52)$$

$$= f_{t+1} (ST_t - ST_{t+1} + SH_{t+1}) - f_t (ST_{t-1} - ST_t + SH_t)$$

Since the grain harvest is realized just once a year, we can treat SH_t to be zero in nonharvest periods or just incorporate this term into ST_t for the period of harvest. Thus, without loss of generality, we can rewrite the demand-for-storage equation (52) as follows:

$$P_{t+1} - P_t = f_{t+1} (ST_t - ST_{t+1}) - f_t (ST_{t-1} - ST_t) \qquad (53)$$

That is, the inter-temporal price difference, $P_{t+1} - P_t$, being defined as the price of storage or the marginal cost of storage, is noted to be a function of changes in stock levels. On the other hand, Brennan [6, p. 56] defines the supply-of-storage equation as the relationship between the price of storage and the current aggregate inventory level only. That is,

$$P_{t+1} - P_t = g_t (ST_t) \qquad (54)$$

Although this hypothesis may be acceptable for seasonally produced commodities harvested over a short time period, Weymar [62, p. 1228] suggests that the supply of storage should be a function of the expected inventory behavior over the time interval between periods t and t+1, especially in explaining the price spread of continuously produced commodities with inventories such as cocoa and pork bellies.

This suggestion is based on the notion that the expected inventory behavior of continuously produced commodities depends upon the quantities produced in each period while the expected inventory level of seasonally produced commodities can reasonably be assumed to decrease from the time of current harvest to the next harvest. Thus, the price spread in the supply-of-storage equation (54) should be dependent upon the expected inventory behavior over the intervening interval.

The main contention of the following discussion, however, is that Weymar's observation is correct but for the wrong reason. The basic departure point of Weymar's from Brennan's argument has arisen from the difference in their fundamental assumptions. While Weymar treats the final stock level, ST_{t+1} to be unknown in period t, Brennan assumes ST_{t+1} to be exogenously determined. However, the assumption of a known ST_{t+1} does not justify the supply of storage to be a function of the current inventory level, ST_t. The following argument shows that the demand-for-storage equation, (53), is a function of the expected inventory behavior, which forces the supply-of-storage equation (54) to be also expressed as a function of the expected inventory behavior for an equilibrium solution to exist.

Assume that we are in period t, and P_t, ST_{t-1}, and ST_t are known in the demand-for-storage equation (53). However, P_{t+1} and ST_{t+1} are not known simply because the future has not yet arrived. Then, a logical conclusion to be drawn about P_{t+1} (or ST_{t+1}) is that P_{t+1} (or ST_{t+1}) is a function of expected market conditions in period t+1. That is, if P_t is a function solely of the change in stock levels, $ST_{t-1} - ST_t$, then the price of period t+1 expected at period t should be a function of the expected change in stock levels from period t to t+1. Thus, the uncertainty of the price spread between two periods is attributable to variations in the stock level expected to exist at the end of the next period.

If we consider a distant future period, say t+k, then we would observe from the equation (53) the following relationship:

$$P_{t+k} - P_t = f_{t+k} (ST_{t+k-1} - ST_{t+k}) - f_t (ST_{t-1} - ST_t) \tag{55}$$

This equation seems to state that the inventory behavior over the intervening interval between periods t+1 and t+k-2 does not play any role in determining the price spread, $P_{t+k} - P_t$, as was noticed by Weymar [62, p. 1226]. Faced with the question of what determines the levels of expected inventories such as ST_{t+k-1} and ST_{t+k} at the present period t, we can turn to the definitional equation (50) for a valuable insight. We can rewrite the equation (50) for the (t+k-1)-th period as

$$ST_{t+k-1} = ST_{t+k-2} + SH_{t+k-1} - C_{t+k-1} - DX_{t+k-1} \tag{56}$$

whose general form can be expressed as

$$ST_{t+k-1} = ST_{t-1} + \sum_{i=0}^{k-1} (SH_{t+i} - C_{t+i} - DX_{t+i}) \tag{57}$$

Thus, we note that the expected level of stocks at the end of the period (t+k-1) is equal to the initial stock minus the expected total consumption and export during the intervening periods if no harvest is assumed. Furthermore, based on the equation (50), the equation (57) can be equivalently written as

$$ST_{t+k-1} = ST_{t-1} + \sum_{i=0}^{k-1} (ST_{t+i} - ST_{t-1+i}) \tag{58}$$

By substituting either the equation (57) or (58) into the equation (55), we find that the expected price spread, $P_{t+k} - P_t$, is a function of the expected inventory behavior over the intervening interval, where the expected inventory behavior is synonymous with the expected consumption and export behavior. Because the demand for storage is thus a function of the expected inventory behavior, so must the supply of storage be a function of the expected inventory behavior in order for an equilibrium price spread to exist. This finding supports Weymar's conclusion that the supply of storage is a function of expected inventory behavior.

For an expository purpose, let us consider the equation (53) and its rationale as presented by Brennan. If the ending stock level, ST_t, is to be increased, then the equations (47), (48), and (50) imply that the positive excess demand for the stocks in storage would induce an increase in P_t, which in turn will cause the consumption at t, $C_t + DX_t$, to be decreased. However, the increase in ST_t would in general put downward pressure on P_{t+1} because more stocks will be carried into the period t+1. Similarly, the decrease in ST_t will bring upward pressure on P_{t+1} while P_t is being depressed in period t. Therefore, we can hypothesize that ST_t and $P_{t+1} - P_t$ are inversely related, where the price spread, $P_{t+1} - P_t$, can be positive or negative.

As for the predetermined stock level of ST_{t-1}, which enters into the period t as a carry-over, we note that ST_{t-1} and P_t will be inversely related because the increase in ST_{t-1} implies a larger available supply in period t and thus, tends to depress the current price, P_t. Thus, ceteris paribus, ST_{t-1} and $P_{t+1} - P_t$ are expected to move in the same direction. We also find a similar result when ST_{t+1} is considered. That is, an increase in ST_{t+1} would reduce the quantity consumed in period t+1, thus inducing an increase in P_{t+1}, while a decrease in ST_{t+1} induces a decrease in P_{t+1}. Therefore, under the ceteris paribus assumption, ST_{t+1} and $P_{t+1} - P_t$ will move together. In other words, we can hypothesize on the basis of the equation (53) that

$$\frac{\partial(P_{t+1}-P_t)}{\partial ST_{t-1}} > 0 \tag{59}$$

$$\frac{\partial(P_{t+1}-P_t)}{\partial ST_t} < 0$$

and

$$\frac{\partial(P_{t+1}-P_t)}{\partial ST_{t+1}} > 0$$

Up to now, however, the importance of the quantity harvested, SH_t, has been largely ignored. The justification for this treatment is based on the characteristic of grain production. That is, in the case of corn, for instance, the corn is harvested only once a year: around September and October. Therefore, we did not bring out the role of SH_t in determining the inter-temporal price difference during our previous discussion. However, we shall now examine the specific role played by SH_t within the demand-for-storage model.

Let us assume that there is no actual harvest in periods t and t+1, but an estimate (or a forecast) of the crop size for some future period t+k is available in the beginning of period t+1. If the total storage space is limited, then the storage decision on how much grain to hold in period t+1 should be revised not because of the actual harvest, but because of the expected new crops. Not only is this type of expectation on the future crop size and other general expectations important, but also the factors affecting the current consumption in period t+1 play an important role in determining the level of stocks at the end of period t+1. Therefore, we can state, on the basis of the equation (51b), that the price in period t+1 is a function of the carry-over stock, the factors affecting the current consumption, and the expectations set, which can be divided into two parts--Q_{t+1} and Ω_{t+1}--where the former represents the size of crop in period t+k expected in period t+1 and the latter represents other relevant expectations. Therefore, the actual price observed in period t+1 can be written as

$$P_{t+1} = f_{t+1} (ST_t - ST_{t+1}) \qquad (60)$$

where

$$ST_{t+1} = g_{t+1}\ (\phi_{t+1},\ Q_{t+1},\ \Omega_{t+1}). \tag{61}$$

g_{t+1} indicates the functional relationship unique to period t+1, and ϕ_{t+1} represents the factors affecting the current consumption such as the number of grain-consuming animals and the export demand in period t+1. By the same token, if we assume t to be the current period instead of t+1, we would expect the price during period t to be

$$P_t = f_t\ (ST_{t-1} - ST_t) = f_t\ (ST_{t-1} - g_t\ (\phi_t,\ Q_t,\ \Omega_t)) \tag{62}$$

However, in the beginning of period t+1, all the periods t, t-1, ..., and t-i are a part of the past whose outcomes can not be altered and must be treated as exogenous by the market participants. The expectations currently formed on the basis of some information such as the coming crop size, export demand, the number of animals on feed, etc. will, however, influence the market trader's decision (which is represented by the current market price) if, and only if, the current expectations are different from the expectations formed prior to the current period. Thus, the actual price change observed between periods t and t+1 should be a function of the changes in the expectation sets, not a function of the expectations themselves. We can state the inter-temporal price spread between periods t and t+1 as

$$P_{t+1} - P_t = f_{t+1}\ (ST_t - ST_{t+1}) - f_t\ (ST_{t-1} - ST_t) \tag{63}$$

$$= f_{t+1} [g_t (\phi_t, Q_t, \Omega_t) - g_{t+1} (\phi_{t+1}, Q_{t+1}, \Omega_{t+1})]$$

$$- f_t (ST_{t-1} - ST_t)$$

$$= H_{t+1} [(ST_{t-1} - ST_t), (\phi_{t+1} - \phi_t), (Q_{t+1} - Q_t),$$

$$(\Omega_{t+1} - \Omega_t)]$$

where H_{t+1} represents a functional relationship in period t+1. Since $(ST_{t-1} - ST_t)$ is known prior to period t+1, the influence of the past variables is felt only when new expectations are being formed. Therefore, equation (63) shows that the inter-temporal price spread between periods t and t+1 depends upon the past consumption, and the changes in expectations on consumption demand (i.e., $\phi_{t+1} - \phi_t$), crop size (i.e., $Q_{t+1} - Q_t$), and other variables (i.e., $\Omega_{t+1} - \Omega_t$) such as the interest rate, government policies, etc.

However, when there are no changes in the expectations, for instance, will the price spread be zero? A theoretical answer to this question is "no" because ever present is the influence of diminishing stock levels from period to period until a new harvest is obtained. These diminutions in stock levels are affected by the factors directly affecting consumption. That is, the larger the number of grain-consuming animals being fed the less will be the stock left in storage at the end of the period, which induces a greater pressure for the price of storage to rise between periods t and t+1 due to increased scarcity of stocks in period t+1. Even though other factors do determine the quantity that has disappeared in the past period t, we can assume

$$ST_t - ST_{t-1} = J_t\ (GCAU_t) \tag{64}$$

where $GCAU_t$ is the number of grain-consuming animal units in period t and J_t indicates a functional relationship in period t. By substituting this relationship (64) into the equation (63) and by assuming that the functional relationship (63) is linear, we can rewrite it as

$$P_{t+1} - P_t = \alpha GCAU_t + \beta(Q_{t+1} - Q_t)$$

$$+ \gamma(\phi_{t+1} - \phi_t) + \delta(\Omega_{t+1} - \Omega_t) \tag{65}$$

where α, β, γ, and δ are coefficients. The period-to-period changes in other consumption demand (i.e., $\phi_{t+1} - \phi_t$) and the period-to-period changes in the expectations held by the market traders (i.e., $Q_{t+1} - Q_t$ and $\Omega_{t+1} - \Omega_t$) can not be treated as constant. However, we note that if the crop size for the (t+k)-th period expected at period t+1 is larger than previously anticipated (i.e., $Q_{t+1} > Q_t$), then there will be an attempt on the inventory holder's part to hold less stocks at the end of period t+1 in expectation of a lower price in period t+k due to new harvest coming into the market. This attempt to lower the inventory level will become more visible as the harvest time approaches. Thus, we note, for the case of a larger expected harvest, all price levels will tend to be lower due to the attempted reduction of the current stock levels (i.e., $\beta < 0$). This tendency will be fully materialized when the actual harvest enters into the market. That is, even though progressively smaller levels of inventory would induce higher price levels as the end of the nonharvest period approaches, the expectation of a new crop will exert a strong downward pressure on the current prices so that the price spread would be negative.[1] This negative spread will, of course, become positive as the influence of the abundant stocks (or the over-supply of a commodity caused by the crop newly harvested) wears out and the scarcity of the commodity once again dominates the market transaction with the passage of time.

As for the changes in consumption demand (i.e., $\phi_{t+1} - \phi_t$), the price spread will widen in a positive direction if the current consumption demand is larger than the consumption demand of the previous period. When the market is bullish because of the increased animal units on feed in period t+1, for instance, the price in period t+1 will rise above the price level obtained in period t. On the other hand, if the market is bearish in period t+1 due to the lack of export demand, then the price spread will be

---

[1]At times, this pressure is real, especially near the harvest. Grain Market News issued on September 11, 1981, describes the August cash corn market as "yellow corn markets declined under pressure of increased farm selling as ... farmers started cleaning out farm storage in preparation for the new crop" [56, p. 2].

narrower than otherwise. Thus, we can hypothesize that γ in equation (65) is positive.

The changes in the general expectations set (i.e., $\Omega_{t+1} - \Omega_t$), however, do not have a predictable a priori influence on the price spread because the composition of Ω_{t+1} or Ω_t can not be treated as fixed during all periods. Thus, for estimation purposes and for the sake of simplicity, we will initially assume that

$$\gamma(\phi_{t+1} - \phi_t) + \delta(\Omega_{t+1} - \Omega_t) = \varepsilon_{t+1} \tag{66}$$

where ε_{t+1} represents all random influences on the price spread between periods t and t+1. By substituting this equation (66) into (65), we thus obtain

$$P_{t+1} - P_t = \alpha GCAU_t + \beta(Q_{t+1} - Q_t) + \varepsilon_{t+1} \tag{67}$$

which states that the price spread between periods t and t+1 is composed of the constant consumption factor which induces the notion of stock scarcity, the change in the crop expectations, and the random influences.

The naive assumption of (66) can be replaced by an alternative assumption that the changes in consumption demand and general expectations sets can be partly captured by the spread movement of the previous periods. That is, we alternatively assume

$$\gamma(\phi_{t+1} - \phi_t) + \delta(\Omega_{t+1} - \Omega_t) = \mu(P_t - P_{t-1}) + \omega_{t+1} \tag{68}$$

where μ is a coefficient and ω_{t+1} is the random disturbance. Since the magnitude and the sign of δ is indeterminate, we note that an *a priori* judgment on μ is not possible. However, if μ is negative, $(\phi_{t+1} - \phi_t)$ and $(\Omega_{t+1} - \Omega_t)$ are assumed to have had a narrowing effect on the present price spread; if μ is positive, their effect is thought to be the opposite; and if μ is zero, no effect is assumed. Thus, by substituting the equation (68) into (65), the estimable equation for the impact of the crop size expectations on the price can be written as

$$P_{t+1} - P_t = \alpha GCAU_t + \beta(Q_{t+1} - Q_t) + \mu(P_t - P_{t-1}) + \omega_{t+1} \qquad (69)$$

Thus, the price spread is assumed to be the result of the continuous disappearance of stocks represented by the number of grain-consuming animal units, the change in the expected crop size, the previous movement of the price spread, and random disturbances not captured by the observable variables in the equation.

Empirical Analysis

Based on the theoretical discussion presented up to now, we shall construct specific empirical models and discuss the appropriate statistical methods by which the impact of the USDA corn crop forecasts on daily cash and futures corn prices can be evaluated. After the data set is chosen, the estimated results will be finally discussed.

Assuming that the period t in equation (65) is the day t, we assume that the USDA corn crop forecast becomes available at the end of the trading day t. Thus, the prices observed prior to and including day t are free from the influence of the newly released USDA corn crop forecast while the prices observed after day t are not. Consequently, the price spread between days t and t+1 should contain the influence of this crop forecast if the influence is sizeable enough. Furthermore, the price observed in day t+k (i.e., k days after the day t) will also retain the influence of the USDA corn crop forecast if the surprise element has not subsided before the day t+k.

We must, however, note that the USDA corn crop forecast can not be, by itself, of value to the decision makers because the information received now should be evaluated in terms of other pertinent information previously available. That is, if the newly forecasted crop size is what the traders were expecting all along, then the traders would not alter their decisions based on this redundant information. Thus, if we assume that the crop size privately expected by the traders during the days prior to the new USDA crop forecast announcement is the same as the most recently available USDA crop forecast, then the impact of the newly announced USDA crop forecast on the market prices can be evaluated by comparing it to the previously known USDA crop forecast. Therefore, if we define Q^i_{t+1} (in lieu of Q_{t+1} in the equation (65)) to be the new USDA corn crop forecast announced at the end of the trading day t but used in trading decisions of the following days, t+1, t+2, ..., and t+k of the month i, then it is assumed that the privately anticipated crop size prior to the release of Q^i_{t+1} is

Q_t^{i-1}, the USDA corn crop forecast announced in the previous month (i-1). After the release of Q_{t+1}^{i}, however, the crop size expected by the traders is assumed to remain as Q_{t+1}^{i}. Therefore, the change in expected crop size between days before and after the crop announcement can be written as $(Q_{t+1}^{i} - Q_t^{i-1})$. Let us define P_{t+k}^{ij} (in lieu of P_t in the equation (65)) as the corn price of the j-th month observed k days before (if k is negative) or k days after (if k is positive) the release of the USDA corn crop forecast in month i. If j equals i, then P_{t+k}^{ij} represents the cash price; otherwise, P_{t+k}^{ij} represents the futures price of the j-th futures contract observed on the day (t+k) in month i.

Given these redefined variables of the price and the crop forecast, we can write equations (67) and (69) as

$$P_{t+k}^{ij} - P_t^{ij} = \alpha_{+k}^{ij}\, GCAU_t + \beta_{+k}^{ij}\, (Q_{t+1}^{i} - Q_t^{i-1}) + \varepsilon_{t+k}^{ij} \tag{70}$$

and

$$\begin{aligned} P_{t+k}^{ij} - P_t^{ij} &= \alpha_{+k}^{ij}\, GCAU_t + \beta_{+k}^{ij}\, (Q_{t+1}^{i} - Q_t^{i-1}) \\ &\quad + \mu_{+k}^{ij}\, (P_t^{ij} - P_{t-1}^{ij}) + \omega_{t+k}^{ij} \end{aligned} \tag{71}$$

where α_{+k}^{ij}, β_{+k}^{ij}, and μ_{+k}^{ij} are coefficients,[1] and ε_{t+k}^{ij} and ω_{t+k}^{ij} are the random disturbance terms. We note that equations (70) and (71) are not suitable for analyzing the effect of the anticipated crop

---

[1]The coefficients in equation (70) are not the same as those in equation (71) when the estimation is carried out. However, for the sake of convenience, the same characters are used in both equations.

size in the minds of traders prior to the release of the USDA corn crop forecast. Thus, we must modify the equation (70) as follows to evaluate this effect of anticipated crop sizes.

$$P_t^{ij} - P_{t-k}^{ij} = \alpha_{-k}^{ij}\ GCAU_t + \beta_{-k}^{ij}\ (Q_{t+1}^{i} - Q_t^{i-1}) + \varepsilon_{t-k}^{ij} \qquad (72)$$

where α_{-k}^{ij} and β_{-k}^{ij} are coefficients and ε_{t-k}^{ij} is the random disturbance term. Thus, this equation (72) assumes that the crop size privately expected at day t-k for the day t is the coming USDA corn crop forecast, Q_{t+1}^{i}, and the crop size expected at day t-k for the day t-k is the previous USDA corn crop forecast, Q_t^{i-1}. Thus, if the traders correctly anticipated the coming crop size on day t-k, then β_{-k}^{ij} will not be zero.

Equation (72) has another interpretation. A few days before each USDA corn crop forecast is released, the Leslie Report[1] contains a proprietary forecast of corn crop. Letting P_{t-k} be the last price observed before the release of the Leslie Report, (72) measures the change in price induced by the Leslie Report before the USDA report is released.

Since there is no USDA corn crop forecast in June, the impact of the USDA corn crop forecast in July on the July prices can not be evaluated by estimating the equations (70) through (72). Therefore, the superscript i goes from 1 for August to 4 for November. While j (=i) goes from 1 to 4 for cash prices, in the analysis of

---

[1]For more information about this report, see [40].

futures prices j=1 signifies the nearest month of the futures contract and j=2 the next nearest month of the futures contract, etc. This point will be clarified in the next sections where we discuss the data set and the estimation methods.

Data set

The data set used in this analysis consists of the monthly USDA corn crop forecasts from July 1961 through November 1980, the numbers of grain-consuming animal units, the daily cash corn prices of the central Iowa market,[1] and the daily futures corn prices of the Chicago Board of Trade during the same period.

The source and availability of the monthly USDA corn crop forecast data are discussed in Chapter II. The USDA corn crop forecasts issued prior to 1961 were estimates of all corn while those issued after 1961 exclude the corn for silage and forage, and thus, pertain only to corn for grain. Therefore, the use of USDA corn crop forecasts issued after (and including) 1961 is believed to be more appropriate for the analysis.

Even though the data on number of the grain-consuming animal units on feed are desired on a monthly basis for the purpose of this analysis, no such data are available. The yearly data reported in Feed: Outlook and Situation [56] are used.

---

[1]There is no single market called the central Iowa market. This is the term used in The Des Moines Register [8] for the purpose of reporting prices quoted by country elevators in areas near Des Moines, Iowa.

The series of daily cash corn prices for the central Iowa market is obtained from the daily newspaper, The Des Moines Register [8], for the period 1961-1980. The daily high and low corn cash prices for 5 trading days before and after the announcement of the USDA corn crop forecast were collected for the months of August through November[1] in each year. The daily average corn cash prices were taken to be the simple average of the daily high and low prices.

The daily corn futures prices used in this analysis are those observed in the Chicago Board of Trade between August, 1961, and November, 1980. Depending upon the month of trade, a different number of futures contracts is traded in each month. Even though there are five basic futures contracts--the March, May, July, September, and December contracts--in any one month, we note that six different monthly futures are being traded in the months between January and May. Further, five contracts can be traded in June through August; eight contracts in September; and seven contracts in October, November, and December. For the purpose of our analysis, however, we shall consider the prices of three futures contracts in August and September and prices of two futures contracts in October

[1]The reason that this analysis is limited to these months is that USDA December corn crop estimates were available around the 20th day of December during the years prior to 1971. However, after (and including) 1971, these estimates are made around the 15th day of January of the following year. Therefore, the price analysis based on these December estimates would be difficult to interpret.

and November. The prices of September, December and March futures contracts observed in August and September and those of December and March futures contracts observed in October and November will be analyzed. As in the case of cash prices, the daily average corn futures prices for each contract were obtained by the simple average of the daily high and low futures prices for each of five trading days before and after the announcement of the USDA corn crop forecast.

Estimation methods

The estimation methods for equations (70), (71), and (72) are very similar because their general structure is noted to be in the form of regressing different dependent variables on identical independent variables. That is, the price spreads between two different periods are regressed on the grain-consuming animal units and the changes of two adjacent USDA corn crop forecasts, and these latter two values do not vary from month to month. We shall choose the equation (70) as an example for the purpose of describing and selecting the most appropriate estimation method.

Given the equation (70), we shall concentrate on the impact analysis of USDA corn crop forecasts on the k-th day's cash price in month i. Therefore, we note that $P^{ij}_{t+k} - P^{ij}_{t}$, $GCAU_t$, and $Q^{i}_{t+1} - Q^{i-1}_{t}$ represent the array of values observed on days t and t+k of the month i over the n(=20) observation years. That is, if we simplify the notations of variables in equation (70) for a given day t+k in month i as

$$Y_{jk} = P^{ij}_{t+k} - P^{ij}_{t} \tag{73a}$$

$$G_k = GCAU_t \tag{73b}$$

$$Q_k = Q^{i}_{t+1} - Q^{i-1}_{t} \tag{73c}$$

$$\varepsilon_{jk} = \varepsilon^{ij}_{t+k} \tag{73d}$$

then the equation (70) for the k-th day after the USDA corn crop forecast announcement in month i can be written as

$$Y_{jk} = G_k\alpha_{jk} + Q_k\beta_{jk} + \varepsilon_{jk} \qquad \text{for } k = 1, 2, \ldots, 5 \tag{74}$$

where

$$\alpha_{jk} = \alpha^{ij}_{+k} \tag{75a}$$

$$\beta_{jk} = \beta^{ij}_{+k} \tag{75b}$$

The dimensions of Y_{jk}, G_k, Q_k, and ε_{jk} are all n by 1 (n=20), and α_{jk} and β_{jk} are scalar coefficients unique to the regression of the price spreads between days t and t+k on G_k and Q_k. We can rewrite equation (74) as

$$Y_{jk} = X_{jk}\,\Pi_{jk} + \varepsilon_{jk} \qquad \text{for } k = 1, 2, \ldots, 5 \tag{76}$$

where

$$X_{jk} = [G_k\ Q_k]_{n\times 2} \tag{77a}$$

$$\Pi_{jk} = \begin{bmatrix} \alpha_{jk} \\ \beta_{jk} \end{bmatrix}_{2\times 1} \tag{77b}$$

If we assume that $\varepsilon_{jk} \sim \text{NID}\ (0,\ \sigma^2_{jk}\ I_{n\times n})$, then ordinary least squares estimation of equation (76) will yield the best linear unbiased estimator of Π_{jk}.

Since there are five equations represented by the equation (74) (or (76)), however, this system of five equations can be simultaneously estimated by the technique of "seemingly unrelated equations (SUE)." This technique of SUE, described below, is nothing more than an application of generalized least squares to a system of equations in order to obtain more efficient estimators than obtained by ordinary least squares. We shall examine whether the technique of SUE should be chosen over ordinary least squares for the purpose of estimating the five equations represented by equation (76).

Going back to the equation (74), we note that the Y_{jk}'s are all column vectors, each of which has n(=20) elements in it. If we have g(=5) Y_{jk}'s, then all g such vectors can be written as one gn(=100) column vector by stacking the vectors. Similarly, all ε_{jk}'s can be written as one vector of length gn(=100), while all Π_{jk}'s can be written as one vector of length 2g(=10). The entire system can thus be written as

$$Y_j = X_j\Pi_j + \varepsilon_j \tag{78}$$

which represents the system

$$\begin{bmatrix} Y_{j1} \\ Y_{j2} \\ Y_{j3} \\ Y_{j4} \\ Y_{j5} \end{bmatrix} = \begin{bmatrix} X_{j1} & 0 & 0 & 0 & 0 \\ 0 & X_{j2} & 0 & 0 & 0 \\ 0 & 0 & X_{j3} & 0 & 0 \\ 0 & 0 & 0 & X_{j4} & 0 \\ 0 & 0 & 0 & 0 & X_{j5} \end{bmatrix} \begin{bmatrix} \Pi_{j1} \\ \Pi_{j2} \\ \Pi_{j3} \\ \Pi_{j4} \\ \Pi_{j5} \end{bmatrix} + \begin{bmatrix} \varepsilon_{j1} \\ \varepsilon_{j2} \\ \varepsilon_{j3} \\ \varepsilon_{j4} \\ \varepsilon_{j5} \end{bmatrix} \tag{79}$$

$$gn\times 1 \qquad\qquad gn\times 2g \qquad\qquad 2g\times 1 \qquad gn\times 1$$

By definition, the variance-covariance matrix for ε_j is

$$\Sigma = E(\varepsilon_j \varepsilon_j') = \begin{bmatrix} E(\varepsilon_{j1}\varepsilon_{j1}') & E(\varepsilon_{j1}\varepsilon_{j2}') & \cdots & E(\varepsilon_{j1}\varepsilon_{j5}') \\ E(\varepsilon_{j2}\varepsilon_{j1}') & E(\varepsilon_{j2}\varepsilon_{j2}') & \cdots & E(\varepsilon_{j2}\varepsilon_{j5}') \\ \vdots & \vdots & & \vdots \\ E(\varepsilon_{j5}\varepsilon_{j1}') & E(\varepsilon_{j5}\varepsilon_{j2}') & \cdots & E(\varepsilon_{j5}\varepsilon_{j5}') \end{bmatrix}_{gn\times gn} \tag{80}$$

Each term in the principal diagonal of Σ is an $n\times n$ variance-covariance matrix. Thus, $E(\varepsilon_{jk}\varepsilon_{jk}')$ is the variance-covariance matrix for the disturbance in the k-th equation. Each off-diagonal term in Σ represents an $n\times n$ matrix whose elements are the contemporaneous and lagged covariances between disturbances from a pair of equations [27, pp. 238-239]. That is, if we write

$$E(\varepsilon_{jk}\varepsilon_{jm}') = \sigma_{km} I_{n\times n} \qquad \text{for } k, m = 1, 2, \ldots, 5 \tag{81}$$

where σ_{km} is the covariance between disturbances from the k-th and the m-th equations, then we have assumed that the disturbance in any single equation is homoscedastic (for $k=m$) and serially noncorrelated. The value of the constant variance can, of course, be different in different equations. By substituting (81) into (80), we obtain

$$\Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{15} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{25} \\ \vdots & \vdots & & \vdots \\ \sigma_{51} & \sigma_{52} & \cdots & \sigma_{55} \end{bmatrix} \otimes I_{n\times n} \tag{82}$$

$$= \Sigma_a \otimes I$$

where Σ_a is a g×g (=5×5) matrix and I is a unit matrix of order n×n.[1]

By applying Aiken's generalized least squares to the equation (78), the best linear unbiased estimator of Π_j, $\hat{\Pi}_j$, is found to be

$$\hat{\Pi}_j = (X_j' \Sigma^{-1} X_j)^{-1} X_j' \Sigma^{-1} Y_j \tag{83}$$

$$= (X_j' (\Sigma_a^{-1} \otimes I) X_j)^{-1} X_j' (\Sigma_a^{-1} \otimes I) Y_j$$

which can be rewritten as

$$\hat{\Pi}_j = \begin{bmatrix} \sigma^{11} X_{j1}' X_{j1} & \sigma^{12} X_{j1}' X_{j2} & \cdots & \sigma^{15} X_{j1}' X_{j5} \\ \sigma^{21} X_{j2}' X_{j1} & \sigma^{22} X_{j2}' X_{j2} & \cdots & \sigma^{25} X_{j2}' X_{j5} \\ \vdots & & & \\ \sigma^{51} X_{j5}' X_{j1} & \sigma^{52} X_{j5}' X_{j2} & \cdots & \sigma^{55} X_{j5}' X_{j5} \end{bmatrix}^{-1} \begin{bmatrix} \sum_{k=1}^{5} \sigma^{1k} X_{j1}' Y_{jk} \\ \sum_{k=1}^{5} \sigma^{2k} X_{j2}' Y_{jk} \\ \vdots \\ \sum_{k=1}^{5} \sigma^{5k} X_{j5}' Y_{jk} \end{bmatrix} \tag{84}$$

[1] The symbol $\otimes$ denotes Kronecker multiplication of matrices. For this operation, see Johnston [27, p. 92].

where σ^{mn} denotes the element of the m-th row and the n-th column in Σ_a^{-1}. The variance-covariance matrix of $\hat{\Pi}_j$ is then expressed as

$$\mathrm{Var}\,(\hat{\Pi}_j) = (X_j' \, \Sigma^{-1} \, X_j)^{-1} = (X_j' \, (\Sigma_a^{-1} \otimes I) \, X_j)^{-1} \tag{85}$$

Given this best linear unbiased estimator of Π_j, $\hat{\Pi}_j$, we note from equations (73b) and (73c) that $X_{j1} = X_{j2} = \ldots = X_{j5}$. Therefore, the equation (84) can be written in terms of X_{j1} alone as

$$\hat{\Pi}_j = (\Sigma_a^{-1} \otimes X_j' \, X_j)^{-1} \, (\Sigma_a^{-1} \otimes X_j' \, Y_j) \tag{86a}$$

$$= \begin{bmatrix} (X_{j1}' \, X_{j1})^{-1} & (X_{j1}' \, Y_{j1}) \\ (X_{j1}' \, X_{j1})^{-1} & (X_{j1}' \, Y_{j2}) \\ \vdots & \\ (X_{j1}' \, X_{j1})^{-1} & (X_{j1}' \, Y_{j5}) \end{bmatrix}$$

If the X_{jk} matrix is the same for each equation, i.e., $X_{j1} = X_{j2} = \ldots = X_{jg}$ as in the case of our example, even if the disturbance terms are correlated between each equation, then, the technique of seemingly unrelated equations reduces to ordinary least squares estimation.[1] Consequently, the appropriate method for the general equations (70) and (72) is the ordinary least squares.

---

[1]This is also true if $\sigma_{km} = 0$ for $k \neq m$. See Johnston [27, p. 240] and Kmenta [30, pp. 520-523].

When the equation (71) is estimated, using the cash price differences, we once again discover that the technique of seemingly unrelated equations is reduced to an ordinary least squares estimation method because the independent variables in the k(=5) equations of a given month i are all identical. However, the use of the seemingly unrelated equations technique is justified for estimating equation (71) for the futures prices because the errors in the equations in any given month are believed to be strongly correlated and the independent variables in the system of equations to be estimated are not identical.

Since the equation (71) contains a new variable, we redefine additional terms in (71) as

$$P_j = P_t^{ij} - P_{t-1}^{ij} \tag{86b}$$

$$\mu_{jk} = \mu_{+k}^{ij} \tag{86c}$$

$$\varepsilon_{jk} = \omega_{t+k}^{ij} \tag{86d}$$

Then, equation (71) can be written in a form similar to equation (74) as

$$Y_{jk} = G_k\alpha_{jk} + Q_k\beta_{jk} + P_j\mu_{jk} + \varepsilon_{jk} \qquad \text{for } k = 1, 2, \ldots, 5 \tag{87}$$

The simplification of this equation in the form of equation (76) yields

$$Y_{jk} = X_{jk}\Pi_{jk} + \varepsilon_{jk} \qquad \text{for } k = 1, 2, \ldots, 5 \tag{88}$$

where

$$X_{jk} = [G_k \; Q_k \; P_j]_{n\times 3} \tag{89a}$$

$$\Pi_{jk} = \begin{bmatrix} \alpha_{jk} \\ \beta_{jk} \\ \mu_{jk} \end{bmatrix}_{3\times 1} \tag{89b}$$

Given these newly defined matrices of X_{jk} and Π_{jk}, we can further simplify the equation (88) in the form of equation (78) to simultaneously estimate g(=k=5) equations. Copying the equation (78) gives

$$Y_j = X_j \Pi_j + \varepsilon_j \tag{90}$$

where Y_j, X_j, Π_j, and ε_j have a dimension of (gn×1), (gn×3g), (3g×1), and (gn×1), respectively.

Suppose, however, that futures prices of h futures contract months are studied in month i. Then, we can construct the following system of equations for j = 1, 2, ... h as

$$Y = X\Pi + \varepsilon \tag{91}$$

which represents the system

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_h \end{bmatrix} = \begin{bmatrix} X_1 & 0 & \dots & 0 \\ 0 & X_2 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & X_h \end{bmatrix} \begin{bmatrix} \Pi_1 \\ \Pi_2 \\ \vdots \\ \Pi_h \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_h \end{bmatrix} \tag{92}$$

and we see from (89a) that each X_j matrix contains a different P_j when there are more than one j's to be considered. Therefore, the use of the SUE technique for estimating the equation (71) is justified.

The difference between (70) and (72) can be interpreted as follows. Equation (72) measures the actual effects of the Leslie Report or the effects of the anticipated USDA report; (70) measures the effect of the actual USDA report. Their sum measures the sum of the effects of the two reports. If we replace k by r in (72), and add (70) and (72), we obtain

$$Y_{jk} + Y_{jr} = X_{jk} (\hat{\Pi}_{jk} + \hat{\Pi}_{jr}) + (\hat{\varepsilon}_{jk} + \hat{\varepsilon}_{jr}) \quad \text{for } k \neq r \tag{93}$$

where

$$Y_{jk} + Y_{jr} = (P^{ij}_{t+k} - P_t) + (P^{ij}_t - P^{ij}_{t-r})$$

$$= P^{ij}_{t+k} - P^{ij}_{t-r} \tag{94}$$

The matrix of variances and covariances of the sum of the coefficients in (93) can be written as

$$\text{Var } (\hat{\Pi}_{jk} + \hat{\Pi}_{jr}) = \text{Var } (\hat{\Pi}_{jk}) + \text{Var } (\hat{\Pi}_{jr}) + 2 \text{ COV } (\hat{\Pi}_{jk}, \hat{\Pi}_{jr}) \tag{95}$$

and

$$\text{COV } (\hat{\Pi}_{jk}, \hat{\Pi}_{jr}) = E[(\hat{\Pi}_{jk} - \Pi_{jk})(\hat{\Pi}_{jr} - \Pi_{jr})'] \tag{96}$$

$$= E[(X'_{jk}X_{jk})^{-1} X'_{jk} \varepsilon_{jk}\varepsilon'_{jr}X_{jr} (X'_{jr}X_{jr})^{-1}]$$

$$= (X'X)^{-1} X'E(\varepsilon_{jk}\varepsilon_{jr}') X (X'X)^{-1} \qquad \text{because } X_{jk} = X_{jr} = X$$

$$= (X'X)^{-1} \hat{\sigma}_{kr}$$

where $\hat{\sigma}_{kr}$ is the estimated covariance between $\hat{\varepsilon}_{jk}$ and $\hat{\varepsilon}_{jr}$. Therefore, equation (95) can be rewritten as

$$\text{Var } (\hat{\Pi}_{jk} + \hat{\Pi}_{jr}) = (X'X)^{-1} (\hat{\sigma}_{kk} + \hat{\sigma}_{rr} + 2\ \hat{\sigma}_{kr}) \qquad (97)$$

When Var $(\hat{\Pi}_{jk} + \hat{\Pi}_{jr})$ is obtained accordingly, we can conduct t-tests for the significance of each coefficient, $\hat{\Pi}_{jk} + \hat{\Pi}_{jr}$. The results obtained, then, will enable us to examine the impact of the anticipated USDA (or actual Leslie Report) and actual USDA corn crop forecast announcement on the price spreads between any two days within the study period.

Estimation results

The impact of the USDA corn crop forecasts on the daily Iowa corn cash prices and the Chicago Board of Trade corn futures prices is estimated and evaluated in this section, first assuming the relationship (67) and then assuming the relationship (69).

The ordinary least squares regression coefficients of the corn cash price spreads on the number of grain-consuming animal units (GCAU) and on the USDA corn crop forecast changes (CROP) are tabulated in Tables 7 through 10. Prices are measured in cents per bushel, GCAU is measured in millions of animal units, and CROP is measured in tens of millions of bushels.

Table 7, for example, presents the estimated coefficients of equations (70) and (72) for the pre-harvest month of August while Table 8 presents the same for September. That is, the impact of the difference between July and August USDA corn crop forecasts on the August corn cash prices is evaluated in Table 7 while the impact of the difference between the August and September USDA corn crop forecasts is analyzed in Table 8. In those and subsequent tables, GCAU represents the number of grain-consuming animal units and CROP represents the change in the USDA crop forecast of one month from that of the previous month. Therefore, the values of coefficients of GCAU represent the scarcity factors and those of CROP represent the impact of the anticipated report such as Leslie Report (for $P_t - P_{t-k}$) and the impact of the actual report (for $P_{t+k} - P_t$).

Since stocks decline rapidly with lapse of time if GCAU is large, the coefficients of GCAU are expected to be positive. Looking at Tables 7 and 8, we find that of the 22 coefficients of GCAU, however, 12 are negative, and the only two coefficients that are significant at the 5 percent level are negative. Most of the coefficients of GCAU in September are negative. This suggests that the pressure to reduce stocks in preparation for a new harvest dominates the market to lower the prices from day to day, especially during the latter part of September. It should be noted, however, that none of these coefficients are significant.[1]

[1]Equations with intercept terms and equations without GCAU were estimated but are not presented. All the estimated intercepts were nonsignificant.

Table 7. Estimated coefficients when the Iowa August corn cash price spreads are regressed on GCAU and the differences of the August and July USDA corn crop forecasts

Dependent variable	August cash GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	-0.0017 (0.0010)[b]	0.0313 (0.0288	0.2309 9.3385	2.10
P_t-P_{t-3}	-0.0019* (0.0008)	0.0206 (0.0215)	0.3521 5.2181	3.80*
P_t-P_{t-2}	-0.0035+ (0.0017)	0.0395 (0.0486)	0.2688 26.6094	2.57
P_t-P_{t-1}	-0.0029* (0.0012)	-0.0107 (0.0332)	0.2987 12.4487	2.98+
$P_{t+1}-P_t$	0.0017+ (0.0009)	-0.1218** (0.0250)	0.6763 7.0526	14.62**
$P_{t+2}-P_t$	0.0026 (0.0014)	-0.0328 (0.0408)	0.2261 18.7939	2.05
$P_{t+3}-P_t$	0.0023 (0.0014)	0.0238 (0.0383)	0.1827 16.5642	1.57
$P_{t+4}-P_t$	0.0027+ (0.0015)	0.0301 (0.0427)	0.1984 20.5844	1.73
$P_{t+5}-P_t$	0.0028 (0.0104)	0.0016 (0.0465)	0.1700 24.3970	1.43
$P_{t+3}-P_{t-4}$	0.0006 (0.0021)	0.0614 (0.0607)		
$P_{t+4}-P_{t-4}$	0.0011 (0.0022)	0.0329 (0.0699)		

[a] R^2 is the coefficient of determination for the regression estimated; S_e^2 is the residual mean square for the regression and is listed below the value of R^2.

[b] The values in the parentheses are the estimates of standard errors.

+Associated with a coefficient indicates that this coefficient is significantly different from zero while + associated with the F values indicates that not all coefficients in the equation are different from zero at the 10% significance level.

*Conveys the same meaning as explained in + above, but the statistical test is conducted at the 5% significance level.

**Conveys the same meaning as explained in + above, but the statistical test is conducted at the 1% significance level.

Table 8. Estimated coefficients when the Iowa September corn cash price spreads are regressed on GCAU and the differences of the September and August USDA corn crop forecasts[1]

Dependent variable	September cash GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	0.0008 (0.0024)[b]	0.1044 (0.1130)	0.0528 64.9991	0.50
P_t-P_{t-3}	0.0001 (0.0017)	0.0985 (0.0796)	0.0794 32.2990	0.78
P_t-P_{t-2}	-0.0005 (0.0009)	0.0179 (0.0434)	0.0266 9.6124	0.25
P_t-P_{t-1}	-0.0002 (0.0007)	-0.0229 (0.0356)	0.0293 6.4442	0.27
$P_{t+1}-P_t$	-0.0000 (0.0010)	0.0007 (0.0490)	0.0000 12.2220	0.00
$P_{t+2}-P_t$	-0.0011 (0.0010)	0.0669 (0.0498)	0.1316 12.6394	1.36
$P_{t+3}-P_t$	-0.0014 (0.0020)	0.0544 (0.0977)	0.0390 48.5697	0.37
$P_{t+4}-P_t$	-0.0005 (0.0028)	0.1014 (0.1312)	0.0329 87.6415	0.31
$P_{t+5}-P_t$	-0.0012 (0.0021)	0.1092 (0.1000)	0.0739 50.9215	0.72
$P_{t+3}-P_{t-4}$	-0.0006 (0.0025)	0.1588 (0.1160)		
$P_{t+4}-P_{t-4}$	0.0003 (0.0033)	0.2058 (0.1576)		

[1]For footnotes, see those in Table 7.

The coefficients of CROP are expected to be negative because the larger the USDA corn crop forecast of this month is, when compared to that of the previous month, the larger the corn marketings during the current and future periods will be. However, the expected relationship between the price spreads and August CROP held only for 3 days around the announcement day and only one of these coefficients (for $P_{t+1}-P_t$) is significant. While only one coefficient of September CROP is negative, it is smaller than its standard error. Since August CROP had a statistically significant impact only for the cash price observed on the day immediately following the announcement day in August, the impact of the August USDA corn crop forecast was not anticipated before its announcement.

We now turn to the price spreads $(P_{t+3}-P_{t-4})$ and $(P_{t+4}-P_{t-4})$ which are tabulated at the bottom two rows of Tables 7 and 8.[1] The derived estimated results of these two price spreads are presented[2] because of the belief that these spreads capture the additive effects of the anticipated USDA report or of the Leslie Report and the actual USDA report. However, the positive coefficients

---

[1]The estimation method for these price spreads from the previously estimated equations was discussed in the later part of the previous section.

[2]Given the 9 original equations estimated, we can compute $9!/(2!7!)$, which equals 36, different price spreads. Since this was not the main objective of this dissertation (and even though the interpretation of these new price spreads may yield interesting insights to the anticipated and actual impact of crop information), other price spreads were not examined.

associated with August and September CROP's show that the hypothesized relationship between the price spread and the crop information did not hold and their values are not statistically significant, which possibly indicates that the impacts of the anticipated and actual crop reports wear out quickly within the period examined.

When the significance of all the coefficients in the estimated equation were tested by an F-test,[1] only two out of 18 F values in Tables 7 and 8 showed that not all coefficients in the equation were zero at the 5 percent significance level. However, only meaningful interpretation could be found in $(P_{t+1}-P_t)$ observed in August. That is, the coefficient of GCAU was positive and that of CROP was negative for $(P_{t+1}-P_t)$ in August, as hypothesized. The significance of this equation thus showed that the actual USDA corn crop forecast did affect the corn price observed on the day immediately following the announcement day in August.

The respective impact of the anticipated and actual October and November USDA corn crop forecasts on the October and November corn cash prices is tabulated in Tables 9 and 10.[2] There, we can

---

[1]An F-test tests the null hypothesis of all coefficients being zero against the alternative hypothesis of not all of them being zero as was shown in (36). The interpretation and estimation of R^2 and F values were also discussed in (33a) and (40). For more information, also see [38, 50, and 63].

[2]The estimated results of $(P_{t+3}-P_{t-4})$ and $(P_{t+4}-P_{t-4})$ are not presented in these tables because there was no significant correlation between the residuals of the estimated equations to be considered.

Table 9. Estimated coefficients when the Iowa October corn cash price spreads are regressed on GCAU and the differences of the October and September USDA corn crop forecasts[1]

Dependent variable	October cash			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	-0.0017 (0.0011)[b]	0.0750 (0.0840)	0.1415 14.5713	1.48
P_t-P_{t-3}	0.0007 (0.0024)	-0.5343** (0.1776)	0.3358 65.1578	4.55*
P_t-P_{t-2}	0.0001 (0.0018)	-0.3444* (0.1331)	0.2712 36.6027	3.35+
P_t-P_{t-1}	0.0002 (0.0018)	-0.3587* (0.1368)	0.2763 38.6957	3.44+
$P_{t+1}-P_t$	-0.0016 (0.0015)	-0.0324 (0.1098)	0.0714 24.8921	0.69
$P_{t+2}-P_t$	-0.0022 (0.0013)	-0.1083 (0.0963)	0.1934 19.1781	2.16
$P_{t+3}-P_t$	-0.0026* (0.0012)	-0.0698 (0.0881)	0.2399 16.0562	2.84+
$P_{t+4}-P_t$	-0.0017 (0.0011)	-0.0351 (0.0836)	0.1286 14.4258	1.33
$P_{t+5}-P_t$	-0.0021+ (0.0012)	0.0936 (0.0910)	0.1847 17.1213	2.04

[1]For footnotes, see those in Table 7.

Table 10. Estimated coefficients when the Iowa November corn cash price spreads are regressed on GCAU and the differences of the November and October USDA corn crop forecasts[1]

Dependent variable	November cash GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	-0.0000 (0.0010)[b]	-0.1265 (0.0761)	0.1485 10.2533	1.57
P_t-P_{t-3}	-0.0005 (0.0011)	-0.0814 (0.0845)	0.0779 12.6533	0.76
P_t-P_{t-2}	0.0002 (0.0007)	-0.1075* (0.0508)	0.2029 4.5724	2.29
P_t-P_{t-1}	-0.0003 (0.0004)	-0.0552+ (0.0282)	0.2488 1.4085	2.98+
$P_{t+1}-P_t$	0.0011 (0.0010)	-0.1285 (0.0807)	0.1349 11.5347	1.40
$P_{t+2}-P_t$	0.0008 (0.0018)	-0.0405 (0.1419)	0.0108 35.6245	0.10
$P_{t+3}-P_t$	0.0012 (0.0018)	-0.0294 (0.1398)	0.0249 34.6011	0.23
$P_{t+4}-P_t$	0.0016 (0.0017)	-0.0004 (0.1300)	0.0517 29.9227	0.49
$P_{t+5}-P_t$	0.0012 (0.0022)	0.0837 (0.1685)	0.0417 50.2323	0.39

[1]For footnotes, see those in Table 7.

render a similar interpretation for the negative coefficients found with GCAU, which are generally observed during the days after the October crop forecast and the days prior to the November crop forecast announcement. That is, the actual harvest does enter into the market to lower the general price levels during October and November. What is interesting in Tables 9 and 10, however, centers on the fact that the hypothesized negative relationship between cash price spreads and CROP (= the difference of two adjacent USDA crop forecasts) has generally prevailed and that a statistically significant impact of the USDA crop forecast is felt during the days prior to its announcement, especially in October. That is, on the third day prior to the October USDA corn crop forecast announcement day, the traders actively discount the prices by anticipating the coming crop forecast. This is true because two USDA corn crop forecasts have been issued and the corn crop is fully matured by then, and the keen traders are aware of the impact of the to-be-realized crop size on the market prices. Thus, the impact of the anticipated crop size dominates the October cash market, and the adjustment is accordingly made during the three days prior to the crop announcement. Therefore, when the crop forecast is actually made by the USDA, it does not have any impact on the market. A ten-million bushel decrease in the October USDA corn crop forecast from the September crop forecast would increase the price spreads between the third, second, and first days prior to the announcement and the announcement day by 0.5343 cents, 0.3444 cents, and 0.3587 cents, respectively. On the second day, prior to the November USDA corn crop forecast announcement day, the traders anticipate and take into account the coming USDA crop forecast information in determining the prices.

When we examine the coefficients of determination R^2, for all the equations tabulated in Tables 7 through 10, we discover that they range from 0 for the day after the September crop forecast to 0.6763 for the day after the August crop forecast. Since R^2 indicates the ratio of explained variation to total variation, the largest R^2 observed on the day after the August crop forecast partially indicates the strong impact of that crop forecast on the prices. That is, the actual impact of USDA corn crop forecasts on the Iowa cash corn prices is most significant in August while the anticipated impact is most significant in October.

Table 11 presents the ordinary least squares (OLS) results from estimating the relationship between the cash prices, GCAU, the differences of two adjacent USDA corn crop forecasts (CROP), and the cash price spreads between the day immediately preceding announcement and the day of announcement (i.e., the variable BEFORE). Since BEFORE is introduced to capture the systematic residual effects, the coefficients associated with it may be thought as trend coefficients. That is, a positive coefficient represents the continuation of the past trend, while a negative coefficient signifies the reversal of the past trend. Since we are now concentrating on the effect of the USDA corn crop forecast on the prices observed after its announcement, we note in Table 11, as in Table 7, that the August

Table 11. OLS regression coefficients, estimating the impact of the crop forecasts on the cash prices observed during pre-harvest months[1]

Dependent variable	August GCAU	August CROP	August BEFORE	August R^2/S_e^2 [a]	August F	September GCAU	September CROP	September BEFORE	September R^2/S_e^2 [a]	September F
$P_{t+1}-P_t$	0.0002	-0.1274**	-0.5207**	0.8312	21.34**	0.0001	0.0128	0.5263	0.1460	0.97
	(0.0008)[b]	(0.0188)	(0.1507)	3.9598		(0.0010)	(0.0471)	(0.3087)	11.0512	
$P_{t+2}-P_t$	-0.0000	-0.0427	-0.9168**	0.6570	8.30*	-0.0011	0.0666	-0.0112	0.1317	0.86
	(0.0012)	(0.0283)	(0.2269)	8.9703		(0.0011)	(0.0518)	(0.3397)	13.3820	
$P_{t+3}-P_t$	0.0005	0.0170	-0.6285*	0.4253	3.21+	-0.0018	0.0185	-1.5647*	0.3512	3.07+
	(0.0014)	(0.0335)	(0.2683)	12.5433		(0.0017)	(0.0835)	(0.5471)	34.7206	
$P_{t+4}-P_t$	0.0013	0.0250	-0.4783	0.3093	1.94	-0.0009	0.0562	-1.9715*	0.3093	2.54+
	(0.0017)	(0.0413)	(0.3311)	19.1009		(0.0024)	(0.1154)	(0.7559)	66.2757	
$P_{t+5}-P_t$	0.0029	0.0108	0.0347	0.1705	0.89	-0.0015	0.0763	-1.4350*	0.3152	2.61+
	(0.0020)	(0.0484)	(0.3882)	26.2576		(0.0019)	(0.0895)	(0.5863)	39.8667	

[1]For footnotes, see those in Table 7.

USDA corn crop forecast significantly affects only the cash price observed on the day immediately following the announcement. A ten-million bushel increase in the August USDA corn crop forecast from the July one causes the price to decrease by about 0.13 cents, while the September USDA corn crop forecast has no significant impact on the September corn cash prices. What is interesting, however, is the significance in the values of R^2 when $(P_t - P_{t-1})$ is added to the equations. The R^2 of 0.8312 observed in the first day following the announcement day in August is significantly larger than the R^2 of 0.6763 observed in Table 7 for the same day. We further note that none of the coefficients of GCAU are significant at the 5 percent level in Tables 7, 8, and 11. With one exception, the only statistically significant coefficients in Table 11 are coefficients of BEFORE. In both August and September, four coefficients of BEFORE are negative, and three of these are significant. While the negative coefficients indicate the reversal of price movement, their statistical significance observed in August is more understandable, in that the influence of the past price movements is felt during the days immediately following the announcement.

While the generally positive coefficients associated with GCAU in August indicate the influence of the decreasing stock levels, this is not the case in September. The heavy pressure of a new harvest coming into the market may have caused the values of coefficients to be negative in September.

When Table 12 is examined, we find no significant impact of

Table 12. OLS regression coefficients, estimating the impact of the crop forecasts on the cash prices observed during post-harvest months[1]

Dependent variable	October					November				
	GCAU	CROP	BEFORE	R^2/S_e^2 [a]	F	GCAU	CROP	BEFORE	R^2/S_e^2 [a]	F
$P_{t+1}-P_t$	-0.0017	-0.0009	0.0877	0.0825	0.51	0.0014	-0.0544	1.3418*	0.3251	2.73$^+$
	(0.0015)[b]	(0.1320)	(0.1934)	26.0414		(0.0010)	(0.0808)	(0.6130)	9.5282	
$P_{t+2}-P_t$	-0.0022$^+$	0.0022	0.3081$^+$	0.3479	3.02$^+$	0.0010	0.0124	0.9584	0.0467	0.28
	(0.0012)	(0.1048)	(0.1535)	16.4164		(0.0019)	(0.1578)	(1.1974)	36.3502	
$P_{t+3}-P_t$	-0.0026*	-0.0126	0.1596	0.2866	2.28	0.0016	0.0462	1.3689	0.0993	0.62
	(0.0012)	(0.1033)	(0.1514)	15.9569		(0.0018)	(0.1523)	(1.1553)	33.8416	
$P_{t+4}-P_t$	-0.0017	-0.0575	-0.0626	0.1378	0.91	0.0018	0.0439	0.8027	0.0805	0.50
	(0.0011)	(0.1005)	(0.1473)	15.1137		(0.0017)	(0.1451)	(1.1008)	30.7220	
$P_{t+5}-P_t$	-0.0021$^+$	0.0314	-0.1734	0.2401	1.79	0.0011	0.0732	-0.1905	0.0426	0.25
	(0.0012)	(0.1063)	(0.1558)	16.8971		(0.0023)	(0.1908)	(1.4477)	53.1330	

[1]For footnotes, see those in Table 7.

the USDA corn crop forecasts on the post harvest months' prices. However, the negative coefficients of October GCAU clearly indicate the inflow of the new corn harvest into the market, while the positive coefficients of November GCAU show the diminished effect of this inflow on the prices. Whereas six of the coefficients of BEFORE are significant at the 5 percent level in Table 11, only one is significant in Table 12, possibly indicating that no short-term trend is visible during the post-harvest months.

When the futures prices were analyzed by estimating equations (70) and (72), ordinary least squares estimation yielded the results tabulated in Tables 13 through 22. Tables 13 through 15 present regression results when the September, December, and March futures corn price spreads observed in August were respectively regressed on GCAU and the differences between the August and July USDA corn crop forecasts (i.e., CROP). There, we find the anticipated and actual August USDA corn crop forecast significantly influence the September, December, and March futures prices. A meaningful relationship, however, seems to exist only on the days after the crop announcement. That is, a ten-million bushel increase in the August USDA corn crop forecast when compared to the July one depresses the September, December, and March futures prices of the day after the announcement day by 0.1226 cents, 0.1288 cents, and 0.1232 cents, respectively. The same change in the crop forecast tends to decrease the September, December, and March futures prices of two days after the announcement day by 0.0875 cents, 0.1018 cents,

Table 13. Estimated coefficients when the September futures corn price spreads in August are regressed on GCAU and the differences of the August and July USDA corn crop forecasts[1]

Dependent variable	September futures in August			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	-0.0000 (0.0011)[b]	0.0018 (0.0317)	0.0003 11.3711	0.00
P_t-P_{t-3}	-0.0002 (0.0008)	0.0392 (0.0235)	0.1720 6.2103	1.45
P_t-P_{t-2}	-0.0006 (0.0005)	0.0622** (0.0155)	0.5663 2.7027	9.14**
P_t-P_{t-1}	-0.0009+ (0.0004)	0.0347* (0.0126)	0.4687 1.7931	6.18*
$P_{t+1}-P_t$	0.0003 (0.0006)	-0.1226** (0.0168)	0.7952 3.1925	27.18**
$P_{t+2}-P_t$	0.0019+ (0.0010)	-0.0875** (0.0278)	0.5137 8.7093	7.40**
$P_{t+3}-P_t$	0.0023 (0.0014)	-0.0257 (0.0393)	0.1914 17.4142	1.66
$P_{t+4}-P_t$	0.0034 (0.0022)	-0.0228 (0.0624)	0.1546 44.0039	1.28
$P_{t+5}-P_t$	0.0039+ (0.0022)	-0.0141 (0.0614)	0.1957 42.5562	1.70
$P_{t+3}-P_{t-4}$	0.0023 (0.0023)	-0.0239 (0.0656)		
$P_{t+4}-P_{t-4}$	0.0034 (0.0027)	-0.0210 (0.0788)		

[1]For footnotes, see those in Table 7.

(1054)

in Table 7.

Table 14. Estimated coefficients when the December futures corn price spreads in August are regressed on GCAU and the differences of the August and July USDA corn crop forecasts[1]

Dependent variable	December futures in August			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	-0.0001 (0.0012)[b]	0.0229 (0.0346)	0.0324 13.5466	0.23
P_t-P_{t-3}	0.0000 (0.0011)	0.0428 (0.0304)	0.1247 10.4250	1.00
P_t-P_{t-2}	-0.0004 (0.0007)	0.0659** (0.0191)	0.4737 4.1354	6.30*
P_t-P_{t-1}	-0.0006 (0.0005)	0.0346* (0.0134)	0.3927 2.0278	4.53*
$P_{t+1}-P_t$	0.0004 (0.0006)	-0.1288** (0.0176)	0.7982 3.4874	27.70**
$P_{t+2}-P_t$	0.0026+ (0.0013)	-0.1018* (0.0368)	0.4730 15.2449	6.28*
$P_{t+3}-P_t$	0.0029 (0.0017)	-0.0300 (0.0473)	0.2085 25.2723	1.84
$P_{t+4}-P_t$	0.0040 (0.0026)	-0.0287 (0.0732)	0.1622 60.4443	1.35
$P_{t+5}-P_t$	0.0044 (0.0026)	-0.0333 (0.0727)	0.1935 59.6119	1.68
$P_{t+3}-P_{t-4}$	0.0028 (0.0027)	-0.0071 (0.0780)		
$P_{t+4}-P_{t-4}$	0.0039 (0.0037)	-0.0058 (0.1054)		

[1]For footnotes, see those in Table 7.

Table 15. Estimated coefficients when the March futures corn price spreads in August are regressed on GCAU and the differences of the August and July USDA corn crop forecasts[1]

Dependent variable	March futures in August			
	GCAU	CROP	R^2/S_e^2 [a]	F
$P_t - P_{t-4}$	0.0002 (0.0014)[b]	0.0119 (0.0390)	0.0071 17.1976	0.05
$P_t - P_{t-3}$	0.0000 (0.0012)	0.0359 (0.0327)	0.0798 12.0617	0.61
$P_t - P_{t-2}$	-0.0005 (0.0006)	0.0589** (0.0183)	0.4461 3.7885	5.64*
$P_t - P_{t-1}$	-0.0006 (0.0004)	0.0326 (0.0123)	0.3967 1.7130	4.60*
$P_{t+1} - P_t$	0.0004 (0.0007)	-0.1232** (0.0206)	0.7255 4.7698	18.50**
$P_{t+2} - P_t$	0.0026* (0.0011)	-0.1018** (0.0310)	0.5576 10.8713	8.82**
$P_{t+3} - P_t$	0.0031+ (0.0015)	-0.0446 (0.0433)	0.2823 21.1988	2.75+
$P_{t+4} - P_t$	0.0044+ (0.0024)	-0.0387 (0.0694)	0.2075 54.3442	1.83
$P_{t+5} - P_t$	0.0047+ (0.0024)	-0.0417 (0.0678)	0.2409 51.7973	2.22
$P_{t+3} - P_{t-4}$	0.0033 (0.0028)	-0.0327 (0.0775)		
$P_{t+4} - P_{t-4}$	0.0046 (0.0038)	-0.0268 (0.1054)		

[1]For footnotes, see those in Table 7.

respectively. The coefficients associated with GCAU generally resemble those in Table 7 where the August cash prices were analyzed. That is, negative values were generally observed prior to the August crop announcement while positive values were associated with GCAU during the days after the crop announcement. Therefore, we may conclude that the August USDA corn crop forecast does affect the futures prices observed within two days after the announcement. In addition, we observe that R^2s are generally higher in Table 13 than in Table 7, indicating that the independent variables, GCAU and CROP, explain a relatively large portion of total variation observed in futures price series. The largest R^2 for the September, December, and March futures prices are 0.7952, 0.7982, and 0.7255, respectively, all of which are observed on the day immediately following the announcement day.

Tables 16 through 18 present the estimated regression coefficients when the price spreads in the same September, December, and March futures observed in September were regressed on GCAU and the differences of the September and August USDA corn crop forecasts (i.e., CROP). The results found in Tables 16 through 18 are, however, quite similar to those found in Table 8 where the September cash prices were analyzed. That is, there seems to be no impact of the USDA corn crop forecast on the cash or futures prices observed in September. Even though the hypothesized relationships seem to exist between the price spreads and GCAU, and between the price spreads and CROP for the days after the

Table 16. Estimated coefficients when the September futures corn price spreads in September are regressed on GCAU and the differences of the September and August USDA corn crop forecasts[1]

Dependent variable	September futures in September			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	0.0024 (0.0019)[b]	0.0470 (0.0918)	0.0949 42.9400	0.94
P_t-P_{t-3}	0.0012 (0.0016)	0.1053 (0.0763)	0.1277 29.6692	1.32
P_t-P_{t-2}	0.0004 (0.0007)	0.0421 (0.0312)	0.1104 4.9422	1.12
P_t-P_{t-1}	0.0035 (0.0040)	0.0888 (0.1885)	0.0548 18.0895	0.52
$P_{t+1}-P_t$	0.0010 (0.0010)	-0.0081 (0.0485)	0.0476 11.9718	0.45
$P_{t+2}-P_t$	-0.0003 (0.0012)	-0.0075 (0.0582)	0.0051 17.2587	0.05
$P_{t+3}-P_t$	-0.0003 (0.0012)	0.0151 (0.0590)	0.0066 17.7298	0.06
$P_{t+4}-P_t$	0.0011 (0.0020)	0.0347 (0.0936)	0.0259 44.6304	0.24
$P_{t+5}-P_t$	0.0004 (0.0021)	0.0290 (0.0996)	0.0074 50.5667	0.07

[1]For footnotes, see those in Table 7.

Table 17. Estimated coefficients when the December futures corn price spreads in September are regressed on GCAU and the differences of the September and August USDA corn crop forecasts[1]

Dependent variable	December futures in September			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	0.0023 (0.0016)[b]	0.0632 (0.0762)	0.1396 29.5394	1.46
P_t-P_{t-3}	0.0010 (0.0013)	0.1263+ (0.0643)	0.2040 21.0397	2.31
P_t-P_{t-2}	0.0005 (0.0005)	0.0468+ (0.0239)	0.2184 2.9094	1.23
P_t-P_{t-1}	0.0000 (0.0003)	0.0209 (0.0134)	0.1201 0.9165	0.42
$P_{t+1}-P_t$	0.0009 (0.0011)	-0.0156 (0.0513)	0.0443 13.4061	0.18
$P_{t+2}-P_t$	0.0002 (0.0013)	-0.0369 (0.0623)	0.0198 19.7952	0.02
$P_{t+3}-P_t$	-0.0001 (0.0013)	-0.0112 (0.0635)	0.0025 20.5463	0.02
$P_{t+4}-P_t$	0.0011 (0.0020	-0.0128 (0.0940)	0.0180 44.9668	0.17
$P_{t+5}-P_t$	0.0004 (0.0020)	-0.0046 (0.0969)	0.0020 47.8488	0.02

[1]For footnotes, see those in Table 7.

Table 18. Estimated coefficients when the March futures corn price spreads in September are regressed on GCAU and the differences of the September and August USDA corn crop forecasts[1]

Dependent variable	March futures in September			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	0.0023 (0.0017)[b]	0.0611 (0.0801)	0.1245 32.6496	1.28
P_t-P_{t-3}	0.0011 (0.0014)	0.1142[+] (0.0649)	0.1800 21.4554	1.98
P_t-P_{t-2}	0.0006 (0.0005)	0.0498[+] (0.0243)	0.2446 3.0111	2.91[+]
P_t-P_{t-1}	-0.0001 (0.0003)	0.0192 (0.0134)	0.1105 0.9142	1.12
$P_{t+1}-P_t$	0.0008 (0.0011)	-0.0084 (0.0541)	0.0283 14.9264	0.26
$P_{t+2}-P_t$	0.0003 (0.0014)	-0.0387 (0.0659)	0.0211 22.1334	0.19
$P_{t+3}-P_t$	-0.0001 (0.0014)	-0.0176 (0.0669)	0.0046 22.7984	0.04
$P_{t+4}-P_t$	0.0013 (0.0020)	0.0000 (0.0950)	0.0248 45.9406	0.23
$P_{t+5}-P_t$	0.0007 (0.0021)	0.0111 (0.0985)	0.0069 49.4489	0.06

[1]For footnotes, see those in Table 7.

announcement, these relationships are most consistently observed only for the day immediately following the announcement day. The correlation coefficients of the 27 equations estimated in Tables 16 through 18 are also nonsignificant, ranging from 0.0020 to 0.2446.

Tables 19 and 20 present the estimated regression coefficients when the December and March futures corn prices observed in October were regressed on GCAU and the differences of the October and September USDA corn crop forecasts (i.e., CROP). Since no statistically significant coefficients were associated with the independent variables, we can, in general, conclude that the October USDA corn crop forecast had no impact on the futures prices observed in October. However, the negative sign associated with GCAU may be interpreted as the influence of new harvest flooding into the market during October; and the negative coefficients associated with CROP, in most cases, further verify that the influence of the new crop information is present but with insignificant statistical importance. The values of R^2 are also minimal in that none exceeded more than 0.2.

When the December and March futures prices were regressed on GCAU and the differences of the November and October USDA corn crop forecasts, there seemed to exist an anticipated impact of the crop forecast on the fifth day prior to the crop announcement day, as shown in Tables 21 and 22. That is, an anticipated ten-million bushel increase in the November USDA corn crop estimate, when compared to the October one, would decrease the December and March futures price spreads between the fifth day prior to announcement

Table 19. Estimated coefficients when the December futures corn price spreads in October are regressed on the differences of the October and September USDA corn crop forecasts[1]

Dependent variable	December futures in October			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	0.0002 (0.0013)[b]	0.1108 (0.0995)	0.0669 20.4640	0.64
P_t-P_{t-3}	0.0001 (0.0009)	-0.0112 (0.0682)	0.0023 9.6027	0.02
P_t-P_{t-2}	-0.0002 (0.0009)	-0.0671 (0.0663)	0.0563 9.0829	0.54
P_t-P_{t-1}	0.0000 (0.0006)	-0.0563 (0.0419)	0.0914 3.6217	0.91
$P_{t+1}-P_t$	-0.0010 (0.0013)	-0.1115 (0.0953)	0.1001 18.7854	1.00
$P_{t+2}-P_t$	-0.0021 (0.0014)	-0.0572 (0.1034)	0.1299 22.1041	1.34
$P_{t+3}-P_t$	-0.0014 (0.0013)	-0.1444 (0.1007)	0.1519 20.9562	1.61
$P_{t+4}-P_t$	-0.0010 (0.0011)	-0.1081 (0.0832)	0.1233 14.3201	1.27
$P_{t+5}-P_t$	-0.0018 (0.0012)	-0.0091 (0.0870)	0.1210 15.6267	1.24
$P_{t+3}-P_{t-4}$	-0.0012 (0.0012)	-0.0336 (0.0950)		
$P_{t+4}-P_{t-4}$	-0.0008 (0.0015)	0.0027 (0.1113)		

[1]For footnotes, see those in Table 7.

Table 20. Estimated coefficients when the March futures corn price spreads in October are regressed on the differences of the October and September USDA corn crop forecasts[1]

Dependent variable	March futures in October			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	0.0003 (0.0012)[b]	0.1006 (0.0907)	0.0673 16.9957	0.65
P_t-P_{t-3}	0.0000 (0.0009)	-0.0259 (0.0664)	0.0084 9.1174	0.08
P_t-P_{t-2}	-0.0002 (0.0009)	-0.0800 (0.0678)	0.0742 9.5028	0.72
P_t-P_{t-1}	-0.0000 (0.0006)	-0.0664 (0.0457)	0.1056 4.3104	1.06
$P_{t+1}-P_t$	-0.0009 (0.0013)	-0.1148 (0.0961)	0.0975 19.0779	0.97
$P_{t+2}-P_t$	-0.0020 (0.0014)	-0.0503 (0.1034)	0.1147 22.0947	1.17
$P_{t+3}-P_t$	-0.0013 (0.0013)	-0.1180 (0.0988)	0.1251 20.1724	1.29
$P_{t+4}-P_t$	-0.0010 (0.0011)	-0.0646 (0.0851)	0.0726 14.9543	0.70
$P_{t+5}-P_t$	-0.0019 (0.0829)	0.0012 (0.0906)	0.1497 16.9708	1.58
$P_{t+3}-P_{t-4}$	-0.0010 (0.0011)	-0.0174 (0.0860)		
$P_{t+4}-P_{t-4}$	-0.0007 (0.0011)	0.0360 (0.0856)		

[1]For footnotes, see those in Table 7.

Table 21. Estimated coefficients when the December futures corn price spreads in November are regressed on GCAU and the differences of the November and October USDA corn crop forecasts[1]

Dependent variable	December futures in November			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	-0.0005 (0.0007) [b]	-0.1593* (0.0557)	0.3825 5.4977	5.57*
P_t-P_{t-3}	-0.0012 (0.0009)	-0.1208 (0.0721)	0.2712 9.1908	3.35+
P_t-P_{t-2}	-0.0002 (0.0008)	-0.1054 (0.0639)	0.1597 7.2241	1.71
P_t-P_{t-1}	0.0079 (0.0057)	-0.6047 (0.4387)	0.1383 34.0640	1.44
$P_{t+1}-P_t$	0.0002 (0.0010)	-0.0642 (0.0760)	0.0383 10.2185	0.36
$P_{t+2}-P_t$	0.0006 (0.0019)	-0.1016 (0.1431)	0.0277 36.2313	0.26
$P_{t+3}-P_t$	0.0007 (0.0021)	-0.0129 (0.1609)	0.0057 45.8366	0.05
$P_{t+4}-P_t$	0.0012 (0.0023)	-0.0299 (0.1746)	0.0144 53.9337	0.13
$P_{t+5}-P_t$	-0.0000 (0.0027)	0.1143 (0.2093)	0.0179 77.5596	0.16

[1]For footnotes, see those in Table 7.

Table 22. Estimated coefficients when the March futures corn price spreads in November are regressed on GCAU and the differences of the November and October USDA corn crop forecasts[1]

Dependent variable	March futures			
	GCAU	CROP	R^2/S_e^2 [a]	F
P_t-P_{t-4}	-0.0007 (0.0007)[b]	-0.1693** (0.0534)	0.4471 5.0533	7.28**
P_t-P_{t-3}	-0.0013 (0.0009)	-0.1257+ (0.0683)	0.3169 8.2539	4.18*
P_t-P_{t-2}	-0.0003 (0.0008)	-0.1059 (0.0637)	0.1703 7.1906	1.85
P_t-P_{t-1}	-0.0003 (0.0005)	-0.0631 (0.0424)	0.1626 3.1753	1.75
$P_{t+1}-P_t$	0.0000 (0.0010)	-0.0536 (0.0755)	0.0296 10.0810	0.27
$P_{t+2}-P_t$	0.0003 (0.0018)	-0.0925 (0.1418)	0.0231 35.6011	0.21
$P_{t+3}-P_t$	0.0008 (0.0020)	-0.0399 (0.1540)	0.0098 41.9714	0.09
$P_{t+4}-P_t$	0.0014 (0.0023)	-0.0556 (0.1737)	0.0211 53.4190	0.19
$P_{t+5}-P_t$	0.0003 (0.0027)	0.0929 (0.2067)	0.0146 75.6025	0.13

[1]For footnotes, see those in Table 7.

and the day of announcement by 0.1593 cents and 0.1693 cents, respectively. R^2's of 0.3825 and 0.4471 for these two futures price spreads were the largest in each respective futures price series.

Now, we turn to the estimated results of equation (71) for the futures price spreads. The technique of seemingly unrelated equations (SUE) was used for estimating (71) for the futures price relationships because the set of independent variables was not identical for each equation in the system, and nonzero correlations between the disturbance terms in various equations were expected. The gain in efficiency from SUE over OLS increases directly with the correlation between the disturbances from the different equations and inversely with the correlation of the independent variables in the different equation [26, p. 173; 27, p. 241; 30, p. 524]. As discussed earlier, 15 equations are estimated for analyzing the impact of the crop forecasts in each month of August and September, and 10 equations are estimated for October and November. The estimated results are tabulated in Tables 23 through 26, and the corresponding variance-covariance matrices are presented in the Appendix from Table 4 to Table 7.

Before interpreting the results found in Tables 23 through 26, however, the notations used for the dependent variables should be explained. Earlier we had used P^{ij}_{t+k} to denote the futures price of the maturity month j observed on the k-th day from the crop announcement day t in month i, and $P_{t+k}-P_t$ to be the futures price spread between days t and t+k. However, we will now represent this futures price spread as iPjk where i and j now take an alphabetical value of A for August, S for September, O for October, N for November, D for December, and M for March, while k still represents a numerical value of 1 through 5. For example, APS1 represents the September futures price spread between the crop announcement day and 1 day after it, observed in the month of August. APD2, on the other hand, represents the December futures price spread between the crop announcement day and 2 days after it, observed in August. By the same token, therefore, NPD1 represents the December futures price spread between the crop announcement day and 1 day after it, observed in November.

Table 23 thus presents the estimated coefficients when the September, December, and March futures price spreads were regressed on GCAU, the differences of the August and July USDA corn crop forecasts (i.e., CROP), and the respective futures price spreads between the crop announcement day and the day before it (i.e., BEFORE), observed in August. There, we observe that the August USDA corn crop forecast has a statistically significant impact on the September and December futures prices but not on the March futures prices. The impact is felt on the day immediately following the announcement day and is lingering on to the next day. Thus, a ten-million bushel increase in the August USDA corn crop forecast when compared to the July forecast seems to decrease the September and December futures price of the immediately following day by 0.1347 cents and 0.1576 cents, respectively, and by 0.0704 cents

Table 23. SUE coefficients, estimating the impact of the August crop forecast on the futures prices observed in August

Dependent variable	GCAU	CROP	BEFORE
APS1	0.0006	-0.1347**	0.3470**
	(0.0006)[a]	(0.0205)	(0.1141)
APS2	0.0014	-0.0704+	0.5347+
	(0.0010)	(0.0344)	(0.2551)
APS3	0.0021	-0.0142	-0.3476*
	(0.0015)	(0.0500)	(0.1325)
APS4	0.0038	-0.0357	0.3918*
	(0.0026)	(0.0784)	(0.1367)
APS5	0.0052*	-0.0557	1.1840
	(0.0024)	(0.0701)	(0.6431)
APD1	0.0008+	-0.1576**	0.7496
	(0.0004)	(0.0143)	(0.1377)
APD2	0.0026+	-0.0994*	-0.1561
	(0.0013)	(0.0451)	(0.0711)
APD3	0.0029	-0.0274	-0.1459
	(0.0017)	(0.0433)	(0.0943)
APD4	0.0049	-0.0734	1.1420
	(0.0028)	(0.0712)	(0.9748)
APD5	0.0054*	-0.0944	1.7645
	(0.0021)	(0.0617)	(0.4670)
APM1	0.0008	-0.1542	0.8884
	(0.0007)	(0.0241)	(0.4294)
APM2	0.0029	-0.1277	0.4976
	(0.0011)	(0.0369)	(0.2477)
APM3	0.0039	-0.0521	0.3611
	(0.0013)	(0.0476)	(0.1576)
APM4	0.0061	-0.1047	1.8432
	(0.0021)	(0.0594)	(0.4776)
APM5	0.0067	-0.1317	1.5798
	(0.0023)	(0.0737)	(0.6669

Weighted MSE for system = 1.0308 with 195 d.f.
Weighted R^2 for system = 0.8618

[a]The values in the parentheses represent the estimates of the standard errors.

+Indicates that the coefficient is significantly different from zero at the 10% significance level.

*Indicates that the coefficient is significantly different from zero at the 5% significance level.

**Indicates that the coefficient is significantly different from zero at the 1% significance level.

and 0.0994 cents on the second day from the announcement day. Thus, the August USDA corn crop forecast does affect the September and December futures prices for two days after the announcement. The coefficient of determination, R^2, for this system of 15 equations is 0.8618, which indicates the relative success achieved by the model.

In Table 24, we find that the September USDA corn crop forecast is not used in the same degree as the August forecast. That is, no statistically significant coefficient is associated with CROP, the difference between the September and August USDA corn crop forecasts. Even though the term, BEFORE, has positive coefficients that are statistically significant for September and December futures prices, the March futures prices have negative coefficients which are significantly different from zero at the 1 percent significance level. Thus, despite the nonsignificant impact of the September USDA corn crop forecast on the prices, there seems to exist a strong dependency between day-to-day future price spreads.

We find a similar result in Table 25, in that the October USDA corn crop forecast does not exhibit a strong influence on the futures prices observed in October. Even though negative coefficients are associated with CROP during four days after the crop announcement, the coefficient becomes positive on the fifth day after the crop announcement day. While BEFORE exhibits statistical importance of varying levels, the coefficients associated with GCAU do not. However, it is interesting to note that the

Table 24. SUE coefficients, estimating the impact of the September crop forecast on the futures prices observed in September[1]

Dependent variable	GCAU	CROP	BEFORE
SPS1	0.0009	-0.0099	0.0195**
	(0.0010)[a]	(0.0498)	(0.0060)
SPS2	-0.0005	-0.0109	0.0382*
	(0.0012)	(0.0590)	(0.0162)
SPS3	-0.0005	0.0100	0.0571**
	(0.0012)	(0.0590)	(0.0166)
SPS4	0.0009	0.0300	0.0529**
	(0.0020)	(0.0953)	(0.0174)
SPS5	0.0002	0.0240	0.0562*
	(0.0021)	(0.1009)	(0.0250)
SPD1	-0.0002	-0.0047	-0.1627
	(0.0004)	(0.0213)	(0.1095)
SPD2	0.0005	-0.0598	1.3286**
	(0.0008)	(0.0385)	(0.1337)
SPD3	-0.0009	-0.0625	1.5332**
	(0.0009)	(0.0420)	(0.1100)
SPD4	-0.0004	0.0167	0.0529
	(0.0013)	(0.0627)	(0.1422)
SPD5	0.0012	0.0243	-0.4888**
	(0.0020)	(0.0938)	(0.1016)
SPM1	0.0006	0.0242	-0.6960**
	(0.0010)	(0.0479)	(0.2408)
SPM2	-0.0001	0.0156	-2.8202**
	(0.0010)	(0.0484)	(0.3687)
SPM3	-0.0004	0.0187	-1.8873**
	(0.0013)	(0.0608)	(0.2562)
SPM4	0.0011	0.0391	-2.0325**
	(0.0019)	(0.0915)	(0.2870)
SPM5	0.0003	0.0632	-2.7108**
	(0.0019)	(0.0911)	(0.3280)

Weighted MSE for system = 1.1411 with 255 d.f.
Weighted R^2 for system = 0.7941

[1]For footnotes, see those in Table 23.

Table 25. SUE coefficients, estimating the impact of the October crop estimate on the futures prices observed in October[1]

Dependent variable	GCAU	CROP	BEFORE
OPD1	-0.0010	-0.0752	0.6453**
	(0.0012)[a]	(0.0934)	(0.2079)
OPD2	-0.0021	-0.0235	0.5985*
	(0.0013)	(0.1011)	(0.2563)
OPD3	-0.0014	-0.0753	1.2249**
	(0.0010)	(0.0797)	(0.2526)
OPD4	-0.0010	-0.0643	0.7766**
	(0.0009)	(0.0711)	(0.2600)
OPD5	-0.0018	0.0406	0.8825**
	(0.0010)	(0.0743)	(0.2820)
OPM1	-0.0008	-0.0709	0.6605**
	(0.0013)	(0.0954)	(0.1951)
OPM2	-0.0019	-0.0205	0.4491+
	(0.0014)	(0.1053)	(0.2481)
OPM3	-0.0013	-0.0541	0.9617**
	(0.0011)	(0.0871)	(0.2572)
OPM4	-0.0010	-0.0226	0.6331*
	(0.0010)	(0.0795)	(0.2612)
OPM5	-0.0018	0.1262	0.6523
	(0.0011)	(0.0875)	(0.2959)

Weighted MSE for system = 1.0970 with 170 d.f.
Weighted R^2 for system = 0.5387

[1]For footnotes, see those in Table 23.

coefficients of GCAU observed in October are negative while those observed in August are all positive and those in September are mixed. This observation may be explained by the fact that the corn stock level is low in August, while it increases somewhat in September; it most certainly increases in October due to the new harvest. Thus, the sign of the coefficients associated with GCAU seems to reflect the state of a given stock level very well. We also note that R^2 for the system in October is 0.5387, which is lower than 0.7941 observed in the September system; the R^2, however, shows significant improvement of the model when compared to the results found in Tables 19 and 20.

Table 26 presents the final results of regressing the December and March futures price spreads on GCAU, the difference between the November and October USDA corn crop forecasts, and the respective futures price spreads between the day of crop announcement and the day immediately preceding it. The results are similar to those in Tables 21 and 22, in that the November USDA corn crop estimate plays no statistically significant role in determining the December and March futures prices. However, the coefficients of CROP all have the expected negative sign. This seems to indicate that there is a relationship between the November crop forecast and the futures prices, although not a strong one. While generally positive coefficients are associated with GCAU, BEFORE exhibits negative coefficients. Thus, there seems to exist the effect of diminishing stock levels in November, while the December and

Table 26. SUE coefficients, estimating the impact of the November crop estimate on the futures prices observed in November[1]

Dependent variable	GCAU	CROP	BEFORE
PAD1	0.0003 (0.0010)[a]	-0.0697 (0.0780)	-0.0090 (0.0214)
PAD2	0.0007 (0.0019)	-0.1079 (0.1468)	-0.0104 (0.0244)
PAD3	0.0007 (0.0021)	-0.0145 (0.1659)	-0.0026 (0.0291)
PAD4	0.0012 (0.0023)	-0.0334 (0.1791)	-0.0058 (0.0345)
PAD5	0.0005 (0.0028)	0.1080 (0.2159)	-0.0105 (0.0510)
PAM1	-0.0002 (0.0009)	-0.1047 (0.0668)	-0.8095** (0.0993)
PAM2	0.0001 (0.0018)	-0.1471 (0.1396)	-0.8648** (0.1341)
PAM3	0.0005 (0.0020)	-0.0913 (0.1533)	-0.8145** (0.1517)
PAM4	0.0010 (0.0022)	-0.1300 (0.1704)	-1.1772** (0.1869)
PAM5	-0.0003 (0.0025)	-0.0222 (0.1931)	-1.8239** (0.2431)

Weighted MSE for system = 1.0938 with 170 d.f.
Weighted R^2 for system = 0.6825

[1]For footnotes, see those in Table 23.

especially March futures prices after the announcement tend to move in the opposite direction from prices before the announcement. The R^2 for this system of 10 equations observed in November, however, is smaller than the R^2's of August and September but larger than that of October system; this R^2 is substantially larger than the individual values of R^2 in Tables 21 and 22.

In summary, we may conclude this section with the following statements. The August USDA corn crop forecast had an impact on the cash prices observed after the crop announcement while there seemed to exist an impact of anticipated October and November USDA corn crop forecasts two or three days prior to the actual crop announcement. However, when the futures prices were analyzed, there seemed to exist impacts of both anticipated and actual August USDA corn crop forecasts on the September, December, and March futures prices. The September and October USDA corn crop forecasts, on the other hand, had no statistically significant impact on the futures market while there seemed to be some anticipated impact on the fifth day prior to the November USDA corn crop forecast. While the fit of the model represented by the equations (70) and (72) was poor as judged by R^2 of each equation, there was a significant improvement when the past price movement was introduced as in equation (71). The estimation of equation (71) for the five days after the crop announcement day by the technique of seemingly unrelated equations for the futures prices, yielded clearer relationships between the USDA corn crop forecasts and the cash and futures corn prices. There existed a definite impact of the USDA corn crop forecast on the August cash price observed on the day immediately following the crop announcement day, while no other monthly USDA crop forecasts had a similar impact on the cash prices. However, the August USDA corn crop forecast significantly influenced the September and December futures prices observed within two days from the announcement day, while the September, October, and November USDA corn crop forecasts did not exhibit a statistically significant impact on the futures prices. Thus, it can be concluded that the USDA corn crop forecast does have an impact on the cash and futures corn prices observed in August while the other USDA corn crop forecasts do not influence prices.

CHAPTER IV. CONCLUSIONS

This study had two main objectives. They were to examine: (a) the accuracy of the USDA corn crop forecasts issued in July through December for the 1930-1977 period; and (b) the effect of these corn crop forecasts upon daily Iowa corn cash and Chicago Board of Trade corn futures prices for the 1961-1980 period.

In order to pursue objective (a), the five-year revised final estimate which should most accurately represent the true crop size was used as the norm of comparing the monthly USDA corn crop forecasts. This in itself differs from other studies [e.g., 1, 9, 20, 40, and 49] and the extensive use of various statistical tools enabled us to evaluate better the question of accuracy in the monthly USDA corn crop forecasts.

In order to pursue objective (b), we studied the impact of the anticipated and actual USDA corn crop forecast on the daily cash and futures corn prices observed between 5 days before and after its announcement in each month from August to November. We first established a theoretical justification and then an empirical estimation for this task, which differs significantly from those earlier studies [e.g., 19 and 40] that examined only the impact of the actual USDA corn crop forecast on the cash prices without any theoretical justification. The detailed summary and conclusions of the present study follows next.

Assuming that the five-year revised final estimate is the true crop size of a given year, the accuracy of the USDA corn crop forecasts was first analyzed by examining the distribution of the forecast errors which were defined as the differences between the monthly USDA corn crop forecasts and their five-year revised final estimates. The use of the Shapiro-Wilk W statistics enabled us to conclude that the forecast errors for each month from July to December were distributed normally.

When the forecast errors were assumed to be independently distributed over the observation years, an F-test based on Hotelling's T^2 statistics was used to test the null hypothesis of the means of the monthly forecast errors all being zeros. Since this null hypothesis was not rejected at the 5 percent significance level, we drew the conclusion that the means of the monthly USDA corn crop forecasts were not different from the five-year revised final estimates. Therefore, the accuracy of the USDA corn crop forecasts was initially established.

The use of a nonparametric L test, however, showed that the null hypothesis of equal means should be rejected at the 0.001 percent significance level in favor of the alternative hypothesis of ordered means. That is, the magnitude of forecast errors was judged to be progressively smaller from July to December, indicating that the accuracy of the USDA corn crop forecasts improved over the reporting months.

The regression of the five-year revised final estimates on each of the monthly USDA corn crop forecasts showed that the monthly forecasts were accurate in the sense that the intercept terms were statistically not different from zero, while the regression coefficients associated with each monthly forecast were not different from one at the 5 percent significance level. When the five-year revised final estimates were regressed on two or more USDA corn crop forecasts, the only statistically significant coefficient was associated with the most recently issued USDA corn crop forecast. While these findings verified the accuracy and the accuracy improvement of the monthly crop forecasts, they could also be used to forecast the coming crop size. It was concluded that in forecasting the true crop size of a given crop year, the most recent month's forecast was its best estimate.

Therefore, we have found that the monthly USDA corn crop forecasts are accurate and their accuracy improves over the reporting months.

The impact of these USDA corn crop forecasts on the cash and futures corn prices was then analyzed on the basis of the supply-of-storage theory. We first noted that there existed a need to distinguish between producers and suppliers, and between consumers and demanders of a storable commodity. The storable characteristic of corn, therefore, provides an explanation of inter-temporal price spread movements by affecting the stock management behavior of the suppliers and demanders. Because of this fact, the supply-of-storage theory was concluded to be a function of the expected inventory behavior which was influenced by the expected crop information.

Therefore, the empirical models under two different assumptions were constructed to measure the impact of the monthly USDA corn crop forecasts on the daily cash and futures corn prices.

The first assumption treated the price spreads between any two days to be a function of the grain-consuming animal units and the differences of the two adjacent USDA corn crop forecasts. When this relationship was estimated by the ordinary least squares estimation technique for the cash and futures corn prices observed five days before and after the day of the USDA corn crop forecast announcement, we found that only the August crop forecasts had statistically significant impact on the prices. Specifically, the affected prices were the cash corn price observed on the day immediately following the announcement day in August and the September and December futures corn prices observed during the two days after the announcement. Even though an impact of the anticipated USDA corn crop forecasts was present in the October and November cash prices, the September, December, and March futures corn prices observed in August were also influenced by the anticipated crop forecast.

The second assumption introduced a past price spread movement to reduce the random disturbance terms in the model. Ordinary least squares was used to estimate the cash price movements; the SUE technique was used to estimate the impact of the USDA corn crop forecasts on the futures prices.

The qualitative results found were not much different from the results obtained under the first assumption, in that only the August USDA corn crop forecast had a statistically significant impact on the August cash prices and the September and December futures prices observed in August. Therefore, we can conclude that the traders' decisions were most strongly affected by the August USDA corn crop forecast.

Having understood the study within the present framework, however, one can not help noticing a few points by which this study can be further extended or better improved. These points are: (a) analyze the impact of the July USDA corn crop forecast on the cash and futures prices observed in July by taking into account the intended acreage to be planted for corn or some other appropriate information; (b) take more explicit account of the stock level changes rather than relying so heavily on the number of grain-consuming animal units to capture the notion of stock scarcity; and (c) extend this study to examine the impact of the USDA corn crop forecasts on all series of futures prices observed in any given month. When these points are further incorporated into the study, the conclusion drawn from it will be more complete.

BIBLIOGRAPHY

1. Baker, John D., Jr.; and Paarlberg, Don. "Outlook Evaluation - Methods and Results." Agricultural Economics Research 4 (October 1952):105-114.

2. Beckmann, Martin J. "On the Determination of Prices in Futures Market." In Patterns of Market Behavior: Essays in Honor of Philip Taft, pp. 3-16. Edited by Michael J. Brennan. Providence: The Brown University Press, 1965.

3. Bonnen, James T. "Improving Information on Agriculture and Rural Life." American Journal of Agricultural Economics 57 (December 1975):753-763.

4. Brainard, Harry G. Economics in Action. New York: The Oxford University Press, 1959.

5. Breimyer, Harold F. "On Price Determination and Aggregate Price Theory." Journal of Farm Economics 39 (August 1957):676-694.

6. Brennan, Michael J. "The Supply of Storage." American Economic Review 48 (March 1958):50-72.

7. Bullock, J. Bruce. "Some Concepts for Measuring the Economic Value of Rural Data." American Journal of Agricultural Economics 63 (May 1981):346-352.

8. "Central Iowa Markets." The Des Moines (Iowa) Register, 1961 through 1980.

9. Clough, Malcolm. "Changes in Corn Acreage and Production After the Early Indications." Agricultural Economic Research 3 (October 1951):140-146.

10. Cole, Charles L. Microeconomics: A Contemporary Approach. New York: Harcourt Brace Jovanovich, Inc., 1973.

11. Dunn, Edgar S., Jr. Social Information Processing and Statistical Systems - Change and Reform. New York: John Wiley and Sons, Inc., 1974.

12. Eisgruber, L. M. "Micro- and Macro-analytic Potential of Agricultural Information System." Journal of Farm Economics 49 (December 1967):1541-1552.

13. Ezekiel, Mordecai. "Statistical Analyses and the 'Laws' of Price." Quarterly Journal of Economics 42 (November 1927):199-227.

14. Fama, Eugene F. "Efficient Capital Markets: A Review of Theory and Empirical Work." Journal of Finance 25 (May 1970):383-417.

15. Foote, Richard J.; Klein, John W.; and Clough, Malcolm. "The Demand and Price Structure for Corn and Total Feed Concentrates." U.S.D.A. Technical Bulletin No. 1061. Washington, D.C.: U.S. Department of Agriculture, 1952.

16. Gastwirth, J. L. "On Probabilistic Models of Consumer Search for Information." Quarterly Journal of Economics 90 (February 1976):38-50.

17. Geary, R. C. "Moments of the Ratio of the Mean Deviation to the Standard Deviation for Normal Samples." Biometrika 28 (December 1936):295-307.

18. Gibbons, Jean Dickinson. Nonparametric Methods for Quantitative Analysis. New York: Holt, Rinehart, and Winston, Inc., 1976.

19. Gorham, Michael. "Public and Private Sector Information in Agricultural Commodity Markets." Economic Review Federal Reserve Bank of San Francisco (Spring 1978):30-38.

20. Gunnelson, G.; Dobson, W. D.; and Pamperin, S. "Analysis of the Accuracy of U.S.D.A. Crop Forecasts." American Journal of Agricultural Economics 54 (November 1972): 639-645.

21. Henderson, James M.; and Quandt, Richard E. Microeconomic Theory: A Mathematical Approach. New York: McGraw-Hill Book Company, 1971.

22. Hicks, John Richard. Causality in Economics. Oxford: Basil Blackwell, Publisher, 1979.

23. Hieronymus, Thomas A. Economics of Futures Trading for Commercial and Personal Profit. New York: Commodity Research Bureau, Inc., 1971.

24. Hoffman, George. "The Effect of Quarterly Livestock Reports on Cattle and Hog Prices." North Central Journal of Agricultural Economics 2 (July 1980):145-150.

25. Houck, James P.; and Pearson, Daniel. "Official Production Estimates for Corn and Soybeans: Preparation and Accuracy." Minnesota Agricultural Economist No. 578. Agricultural Extension Service, University of Minnesota. April 1976.

26. Intriligator, Michael D. Econometric Models, Techniques, and Applications. Englewood Cliffs: Prentice-Hall, Inc., 1978.

27. Johnston, J. Econometric Methods. New York: McGraw-Hill Book Company, 1972.

28. Kamerchen, David R.; and Valentine, Lloyd M. Intermediate Microeconomic Theory. West Chicago: South-Western Publishing Company, 1977.

29. Kantor, Brian. "Rational Expectations and Economic Thought." Journal of Economic Literature 17 (December 1979):1422-1441.

30. Kmenta, Jan. Elements of Econometrics. New York: Macmillan Publishing Company, 1971.

31. Maddala, G. S. Econometrics. New York: McGraw-Hill Book Company, 1977.

32. Mansfield, Edwin. Microeconomics: Theory and Applications. New York: W. W. Norton and Company, Inc., 1970.

33. Miller, Steve. "The Response of Futures Prices to New Market Information: The Case of Live Hogs." Southern Journal of Agricultural Economics 11 (July 1979):67-70.

34. Milner, Arthur Ross. Grain Marketing: Pricing, Transporting. Westerville: West-Camp Press, 1970.

35. Morrison, Donald E. Multivariate Statistical Methods. New York: McGraw-Hill Book Company, 1976.

36. Muhm, Don. "2nd Largest Corn Crop Seen in Iowa." The Des Moines (Iowa) Register. July 12, 1980:1A.

37. Muth, John F. "Rational Expectations and the Theory of Price Movements." Econometrica 29 (July 1961):315-335.

38. Ostle, Bernard; and Mensing, Richard W. Statistics in Research. Ames: The Iowa State University Press, 1975.

39. Page, Ellis Batten. "Ordered Hypotheses for Multiple Treatments: A Significance Test for Linear Ranks." Journal of American Statistical Association 58 (March 1963):216-230.

40. Pearson, Daniel; and Houck, James P. "Price Impacts of SRS Crop Production Reports: Corn, Soybeans, and Wheat." Unpublished manuscript. Department of Agricultural and Applied Economics, University of Minnesota. April 1977.

41. Pesek, Boris P. "Producer's Quasi-Supply During Stock Period." Journal of Farm Economics 41 (February 1959):103-113.

42. Pliska, Stanley R. "Supply of Storage Theory and Commodity Equilibrium Prices with Stochastic Production." American Journal of Agricultural Economics 55 (November 1973): 653-658.

43. Raikes, Ronald; and Vollink, William. "Measuring Impacts on Demand of Agricultural Commodity Promotion." Southern Journal of Agricultural Economics 7 (December 1975): 161-166.

44. Reid, Frank. "Control and Decontrol of Wages in the United States: An Empirical Analysis." American Economic Review 71 (March 1981):108-120.

45. Rothschild, M. "Searching for the Lowest Price When the Distribution of Prices Unknown." Journal of Political Economy 82 (July/August 1974):689-711.

46. Santomero, Anthony; and Seater, John J. "Inflation-Unemployment Trade-off: A Critique of the Literature." Journal of Economic Literature 16 (June 1978):499-544.

47. Sarle, Charles F. "Adequacy and Reliability of Crop-Yield Estimates." U.S.D.A. Technical Bulletin No. 311. Washington, D.C.: U.S. Department of Agriculture, June 1932.

48. Shapiro, S. S.; and Wilk, M. B. "An Analysis of Variance Test for Normality (Complete Samples)." Biometrika 52 (December 1965):591-611.

49. Smith, Lawrence N. "An Evaluation of Improved Soybean Crop Information." Unpublished Ph.D. Dissertation. Library, Purdue University, 1973.

50. Snedecor, George W.; and Cochran, William G. Statistical Methods. Ames: The Iowa State University Press, 1967.

51. Stein, Jerome L. "The Simultaneous Determination of Spot and Futures Prices." American Economic Review 51 (December 1961):1012-1025.

52. Stigler, George J. "The Economics of Information." Journal of Political Economy 69 (June 1961):213-225.

53. Telser, Lester G. "Futures Trading and the Storage of Cotton and Wheat." Journal of Political Economy 66 (June 1958): 233-255.

54. Thompson, James M. "Analysis of the Accuracy of USDA Hog Farrowing Statistics." American Journal of Agricultural Economics 56 (December 1974):1213-1217.

55. Tomek, William G.; and Robinson, Kenneth L. Agricultural Product Prices. Ithaca: Cornell University Press, 1981.

56. U.S. Department of Agriculture. Agricultural Marketing Service. Grain Market News. Washington, D.C.: U.S. Government Printing Office.

57. U.S. Department of Agriculture. Crop Reporting Board. Preparing Crop and Livestock Estimates. Washington, D.C.: U.S. Government Printing Office, May 1967.

58. U.S. Department of Agriculture. Crop Reporting Board. Crop Reporting Board Catalog: 1981 Release. Washington, D.C.: U.S. Government Printing Office, December 1980.

59. U.S. Department of Agriculture. Economic Research Service. Feed: Outlook and Situation. Fds-282. Washington, D.C.: U.S. Department of Agriculture, August 1981.

60. Waugh, Frederick W. "Demand and Price Analysis: Some Examples from Agriculture." U.S.D.A. Technical Bulletin No. 1316. Washington, D.C.: U.S. Department of Agriculture, November 1964.

61. Wells, Gary J.; and Pittman, Jerold F. "Assessment of the Price Impact of the South Carolina Cucumber Marketing Order." Southern Journal of Agricultural Economics 12 (December 1980):15-18.

62. Weymar, Helmut F. "The Supply of Storage Revisited." American Economic Review 56 (December 1966):1226-1234.

63. Wonnacott, Thomas H.; and Wonnacott, Ronald J. Introductory Statistics for Business and Economics. New York: John Wiley and Sons, Inc., 1972.

64. Working, Holbrook. "The Theory of the Price of Storage." American Economic Review 39 (December 1949):1254-1262.

65. Working, Holbrook. "Hedging Reconsidered." Journal of Farm Economics 35 (November 1953):543-561.

66. Zaltman, Gerald; and Burger, Philip C. Marketing Research: Fundamentals and Dynamics. Hinsdale: Dryden Press, 1975.

ACKNOWLEDGMENTS

As one plods through the rugged road of daily life, many people render their valuable time, effort, and energy to aid him to plod with more comfort. This has truly been the case in my life, and I would like to acknowledge them (of course, not all of them) here with a most sincere gratitude. They are:

My parents, Mr. and Mrs. Chang Kyung Choi,
My parents-in-law, Mr. and Mrs. Jhoon Sik Kim,
My American parents, Mr. and Mrs. R. L. Langlas,
The members of my program-of-study:
Dr. Michael Boehlje, Dr. Dennis Starleaf,
Dr. Gene Futrell, and Dr. Roy Hickman.

I deeply appreciate my American brother, Jim Langlas, for his reading of this dissertation. Mrs. Shirley Baum deserves a special thank-you for her excellent typing of this dissertation and my creative component for a master's requirement. I also thank Dr. Kun Chol Lee for his assistance during the final stage of my graduation.

To two very special people in my life, however, I owe much of what I am. They are my major professor, Dr. George W. Ladd, and my dear wife, Hee Young Kim Choi. Dr. Ladd has patiently encouraged and guided me during my graduate study and especially during my writing of this dissertation. It is he who made me realize the difference between magnificient failure and moderate success in research.

My dear wife, Hee Young, who has given me the joy of being a father to two children--Juno and Juliana--has sacrificed too much too unselfishly throughout my graduate study. Without her sacrifice and encouragement, this dissertation might have been started but never would have been completed.

UMI

Appendix table 2. The coded regression sums of squares from the models in H and H of the

Month	Source	d.f.	Estimate[a]	Month	Source	d.f.	Estimate[a]

Appendix table 1. Necessary estimates to compute the F statistics in the equation (29) (MF_{my}'s are measured in 1,000 bushels)

		Q				
Month	n	A_m	b_m	Mean of MF_{my}	$\sum_{y=1}^{n} (MF_{my})^2$ [a]	S_e^2 [a]
July	44	-23,845.2	0.9998	3,439,607.5	56,687.6437	10.3556859
Aug.	48	-15,014.7	1.0149	3,534,003.3	65,953.7461	4.9497423
Sept.	48	-58,657.8	1.0296	3,533,370.0	65,988.4981	2.3729411
Oct.	48	-99,564.1	1.0354	3,552,310.9	66,597.1597	1.7149441
Nov.	48	-92,710.0	1.0265	3,575,209.6	67,505.5237	1.0181879
Dec.	48	-64,055.2	1.0172	3,579,300.3	67,762.6183	0.5583351

[a]These estimates should be decoded by multiplying them by 10^{10}.

Appendix table 2. The coded regression sums of squares from the models in H_o and H_a of the hypothesis (37) and the residual sum of squares from the model in H_a

Month	Source	d.f.	Estimate[a]	Month	Source	d.f.	Estimate[a]
July	$RSS(\gamma_1)$	1	559.494921	Oct.	$RSS(\gamma_4)$	1	563.089381
	$RSS(\gamma_1)$	1	559.494921		$RSS(\gamma_1,\gamma_2,\gamma_3,\gamma_4)$	4	563.180027
	ESS	43	4.351432		ESS	40	0.666327
	F		0		F		1.8138
Aug.	$RSS(\gamma_2)$	1	561.766611	Nov.	$RSS(\gamma_5)$	1	563.386612
	$RSS(\gamma_1,\gamma_2)$	2	561.767103		$RSS(\gamma_1,\gamma_2,\gamma_3,\gamma_4,\gamma_5)$	5	563.469043
	ESS	42	2.079250		ESS	39	0.377310
	F		0.0099		F		2.1301
Sept.	$RSS(\gamma_3)$	1	562.836771	Dec.	$RSS(\gamma_6)$	1	563.596242
	$RSS(\gamma_1,\gamma_2,\gamma_3)$	3	562.858885		$RSS(\gamma_1,\gamma_2,\cdots,\gamma_6)$	6	563.642076
	ESS	41	0.987468		ESS	38	0.204278
	F		0.4591		F		1.7052

[a]All estimates except the values of F statistic should be decoded by multiplying them by 10^{12}.

Appendix table 3. The matrix of variance-covariance across equations for estimating the coefficients in Table 17[a]

Dependent variable	SPS1	SPS2	SPS3	SPS4	SPS5	SPD1	SPD2
SPS1	10.74						
SPS2	9.82	15.03					
SPS3	4.16	7.15	15.05				
SPS4	2.83	2.21	20.98	39.26			
SPS5	3.86	4.28	21.90	40.32	44.05		
SPD1	0.80	0.46	-0.14	-0.96	-1.88	1.90	
SPD2	4.48	4.05	3.52	4.42	2.77	1.88	6.30
SPD3	3.15	6.90	5.98	5.37	4.94	0.58	4.70
SPD4	4.45	6.24	14.82	22.30	22.51	0.54	5.21
SPD5	6.48	4.90	20.48	37.24	37.90	0.16	6.68
SPM1	9.16	7.44	3.28	2.03	1.88	1.58	4.23
SPM2	6.99	9.64	5.18	0.88	0.88	1.39	3.26
SPM3	2.31	3.40	14.18	21.36	20.87	0.39	3.63
SPM4	2.05	-0.88	18.32	35.42	34.55	-0.04	4.67
SPM5	4.39	2.19	19.29	34.26	34.96	0.62	4.22

[a]Upper triangle of the symmetric variance-covariance matrix is not presented.

SPD3	SPD4	SPD5	SPM1	SPM2	SPM3	SPM4	SPM5
7.57							
7.52	16.88						
6.82	23.16	37.99					
1.52	3.23	5.45	9.85				
3.90	4.41	3.38	7.79	9.94			
4.83	14.67	21.03	3.21	4.98	15.90		
3.52	20.07	34.54	3.74	2.08	21.54	36.07	
3.21	20.34	34.13	5.72	4.38	21.65	34.44	35.76

Appendix table 4. The matrix of variance-covariance across equations for estimating the coefficients in Table 18[a]

Dependent variable	OPD1	OPD2	OPD3	OPD4	OPD5
OPD1	15.08				
OPD2	14.09	17.60			
OPD3	10.07	11.86	10.80		
OPD4	7.41	8.71	9.00	8.49	
OPD5	9.29	10.74	8.37	7.14	9.24
OPM1	15.23	15.33	10.26	7.45	9.34
OPM2	15.09	18.07	11.96	8.74	10.88
OPM3	10.70	13.10	11.35	9.45	9.10
OPM4	8.71	10.71	9.89	9.00	8.23
OPM5	10.83	12.56	9.71	8.32	9.93
	OPM1	OPM2	OPM3	OPM4	OPM5
OPM1	15.69				
OPM2	15.84	19.00			
OPM3	11.29	13.75	12.80		
OPM4	9.16	11.30	11.25	10.56	
OPM5	11.44	13.40	11.57	10.50	12.77

[a]Upper triangle of the symmetric variance-covariance matrix is not presented.

Appendix table 5. The matrix of variance-covariance across equations for estimating the coefficients in Table 19[a]

Dependent variable	NPD1	NPD2	NPD3	NPD4	NPD5
NPD1	8.90				
NPD2	15.74	32.12			
NPD3	17.67	34.75	40.93		
NPD4	18.66	37.48	42.73	47.62	
NPD5	20.92	42.95	49.44	55.02	68.70
NPM1	6.55	13.24	14.32	14.85	15.31
NPM2	13.13	29.08	31.32	33.31	36.89
NPM3	14.49	30.48	35.71	37.05	40.25
NPM4	15.25	33.71	38.08	42.38	46.56
NPM5	15.53	36.62	41.95	46.60	55.35
	NPM1	**NPM2**	**NPM3**	**NPM4**	**NPM5**
NPM1	6.65				
NPM2	13.32	29.20			
NPM3	14.85	30.69	35.21		
NPM4	15.13	33.78	37.71	43.50	
NPM5	15.48	36.98	40.63	47.31	55.72

[a]Upper triangle of the symmetric variance-covariance matrix is not presented.